Breaking Bad: Parameter Uncertainty Caused by Structural Breaks in Stocks

This version: December, 2025

Abstract

We propose "break-age", defined as the time elapsed since the most recent structural break in a stock's return-generating process, as a novel proxy for parameter uncertainty. Unlike existing measures that confound uncertainty with risk, break-age captures the scarcity of relevant historical data caused by regime changes. Using unsupervised machine learning techniques for out-of-sample detection of structural breaks, we show that they are frequent and generate distinct cycles of parameter uncertainty. Stocks with recent breaks exhibit significant negative abnormal returns while the new regime remains ambiguous. Conversely, the statistical resolution of this uncertainty triggers a sharp positive revaluation. Our findings demonstrate that the stability of the data-generating process is a priced and confirm that investors demand a temporary premium for the ambiguity caused by data obsolescence.

Keywords: Stock-age, structural breaks, parameter uncertainty

JEL classification: G12, G14, G17

1 Introduction

Financial economics has long recognized that investors face uncertainty about the true parameters governing the return-generating process, yet empirical measurement of this uncertainty at the stock level remains elusive. Existing proxies such as idiosyncratic volatility (Ang et al., 2006, 2009) or analyst forecast dispersion (Diether et al., 2002) often confound uncertainty with risk and sentiment and therefore provide imperfect signals of the uncertainty investors face. This challenge is particularly acute in light of a large literature showing that financial time series are characterized by structural instability (Pástor and Stambaugh, 2001; Kim et al., 2005; Smith and Timmermann, 2021). Breaks in means, variances, factor premia, and predictive relations occur repeatedly in equity markets and materially alter the parameters central to both asset-pricing inference and portfolio choice (Dangl and Halling, 2012; Smith and Timmermann, 2022; Garlappi et al., 2007). When such breaks occur, pre-break observations become unreliable, forcing investors to learn about a new regime from limited data. The resulting scarcity of relevant information creates precisely the form of parameter uncertainty emphasized in theories of ambiguity aversion and learning: investors temporarily distrust their estimates, reduce exposure to assets with uncertain parameters, and demand higher returns until uncertainty resolves(Pástor, 2000; Pástor and Veronesi, 2009; Collin-Dufresne et al., 2016).

We argue that the time elapsed since the most recent structural break (i.e. "breakage") provides a conceptually clean and economically grounded proxy for this uncertainty. Break-age captures the effective sample size available for learning about the current regime and directly reflects the obsolescence of historical data when the underlying return process shifts. Unlike standard proxies derived from dispersion, volatility, or noise, break-age isolates the uncertainty channel predicted by learning-based and ambiguity-averse frameworks by connecting uncertainty to the stability of the data-generating process itself, rather than to the distribution of realized returns or the presence of trading frictions.

To operationalize this idea, we detect structural breaks at the individual-stock level using a suite of online breakpoint detection methods that mirror the real-time information available to investors, mimicking an out-of-sample setting. We calibrate these methods through extensive simulations to achieve a conservative balance between detection power and false positives, ensuring that identified breaks correspond to economically meaningful regime shifts rather than statistical artifacts. These detected breaks form the basis for an event-study framework in which we examine returns before, during, and after structural breaks, paired with Mahalanobis-distance matching to construct rigorous counterfactual benchmarks that isolate the market response to break-induced uncertainty.

Our analysis reveals a coherent pattern of return dynamics that aligns tightly with the theoretical predictions of learning and ambiguity aversion. Structural breaks are frequent and correspond to substantial shifts in return parameters. At the break-date, stocks experience immediate negative abnormal returns, consistent with a sudden rise in parameter uncertainty. This uncertainty remains elevated between the break and its statistical detection, generating continued return discounts while the new regime remains ambiguous. Once the break becomes statistically identifiable, uncertainty resolves sharply and is accompanied by a pronounced positive revaluation, reflecting the collapse of ambiguity rather than compensation for conventional risk. Across detection methods, this "life-cycle" from uncertainty shock to gradual learning to resolution supports break-age as a powerful and economically meaningful proxy for parameter uncertainty in the cross-section of stock returns.

The remainder of the paper proceeds as follows. Section 2 situates break-age within the broader literature on structural breaks, parameter uncertainty, ambiguity aversion, and learning. Section 3 describes the data, breakpoint detection methods, simulation-based calibration, event-study design, and matching procedure. Section 4 presents the empirical results on the frequency and magnitude of breaks, market reactions, and the dynamics of uncertainty and its resolution. Section 5 concludes.

2 Literature Review

Our research integrates insights from four distinct strands of the asset pricing literature to establish the theoretical and empirical foundations for break-age. We begin by reviewing the evidence on structural breaks in financial time-series, documenting their prevalence and economic significance. Next, we examine the link between parameter instability and ambiguity aversion, outlining how uncertainty takes the form of a priced state variable in equilibrium. We then connect these findings to learning-based asset pricing models, which provide the structural mechanism explaining why the time elapsed since a regime shift governs the precision of investor beliefs. Finally, we evaluate existing proxies for parameter uncertainty, highlighting their susceptibility to confounding risks and motivating the need for a cleaner, history-dependent measure.

2.1 Structural Breaks in Asset Returns

A substantial body of evidence shows that financial data, and asset returns in particular, are subject to structural breaks. Early work, for example, shows that statistical features and financial variables exhibit considerable variation over time. To this extent, Barsky (1989) mentions drastic changes in the correlation between stock

and bond while Dimson et al. (2002) show that long-run correlation patterns of economic variables differ substantially across decades. These findings underscore the time-varying nature of fundamental financial and economic data that is commonly used for applications in asset-pricing or in the context of asset allocation tasks.

The research related to the equity premium provides some of the clearest evidence of breaks in time-series of returns. Pástor and Stambaugh (2001) for example, document several economically large shifts in the U.S. equity premium, stressing that inference about expected returns is highly sensitive to whether such structural breaks are taken into account or not. Similar to Pástor and Stambaugh (2001), Kim et al. (2005) also discover structural breaks in the mean of aggregate stock returns. More recently, Smith and Timmermann (2022) show that these structural breaks extend to cross-sectional risk premia such as investment, size or value which all have significantly decreased to the point of being statistically indistinguishable from 0. Beyond returns, Paye and Timmermann (2006) find that the coefficients of commonly used predictive variables such as the term spread or the dividend yield are also subject to structural breaks. Taken together, these findings highlight that structural breaks are a pervasive feature of the broader financial environment rather than an artifact of a single series or a statistical curiosity.

Against this backdrop, we explore the direct consequences of structural breaks from the perspective of statistical inference and investment decision making. A growing body of literature shows that ignoring structural breaks impedes the validity of predictive relationships (Bossaerts and Hillion, 1999) and limits the power of statistical models (Cremers, 2002). Furthermore, Dangl and Halling (2012) show how structural breaks render predictive relations unstable while Smith and Timmermann (2021) illustrate how they contribute to the difficulty of detecting long-run risk premia. Pesaran et al. (2006) illustrate that the possibility and presence of multiple structural breaks substantially increases estimation risk for means, variances and betas in a prediction framework. From a portfolio selection perspective, unstable parameters undermine the classic mean-variance optimization as outlined by Markowitz (1952). As Kandel and Stambaugh (1996), Barberis (2000), Xia (2001) and Garlappi et al. (2007) illustrate how uncertainty about return parameters directly affects optimal asset allocation. DeMiguel et al. (2009) further demonstrate in an empirical exercise that estimation errors induced by parameter uncertainty is large enough for simple diversification rules (i.e. equally weighting all assets) to outperform sophisticated optimizers in realistic out-of-sample settings. Adding to that, Pettenuzzo and Timmermann (2011) present evidence that suggests that structural breaks in predictive relations create larger distortions for investors than standard estimation uncertainty, underscoring the economic relevance of parameter instability.

4

Given the widespread occurrence and consequences of breaks, recent research has increasingly focused on identifying their economic origins. Ang and Timmermann (2012) link structural breaks in the equity premium to changes in the underlying macroeconomic and policy environment, while Smith and Timmermann (2021) demonstrate that structural breaks in return predictability are connected to changes in the cash-flow growth and discount-rate of companies. In addition to that, Lleo et al. (2020) uses an event-study approach to explain the occurrence of several structural breaks in the variance of Apple returns and discovers that these breaks often happen in the aftermath of a corporate event such as earnings releases. These studies suggest that structural breaks reflect genuine shifts in the economic environment, or are linked to meaningful changes within a company and are therefore likely to alter investors' beliefs and induce periods of heightened parameter uncertainty.

Taken together, the literature establishes two key points. First, structural breaks in financial data are common, economically meaningful and present in various financial time-series. Second, these breaks have significant implications for inference and portfolio choice precisely because they create uncertainty about the true parameters governing the return-generating process. As a consequence of these observations, we argue that if structural breaks occur at the level of individual stocks, then the time since the most recent break, so-called "break-age" naturally captures the local amount of parameter uncertainty investors face. Understanding how they respond to this break-induced uncertainty is therefore central to assessing whether structural instability is priced in the cross-section of returns.

2.2 Parameter Uncertainty, Ambiguity Aversion, and Pricing Implications

Having established that structural breaks are common in financial data and that they instability in the parameters that govern the return process, an important question arises: How does the resulting parameter uncertainty affect investor behavior and asset prices? Before answering this question, we first need to define uncertainty. From an economic perspective, Knight (1921) distinguishes between uncertainty and risk by defining a situation with risk as one in which outcomes are unknown but follow a known probability distribution. Contrary to that, uncertainty refers to a situation in which neither the outcomes nor the probability distribution of outcomes is known (i.e. so-called "Knightian Uncertainty"). Based on this definition, Ellsberg (1961) illustrates that economic agents systematically prefer risk (i.e. gambles with known probabilities) over uncertainty (i.e. gambles with unknown probabilities). This preference for certainty forms the conceptual foundation for understanding investor responses to parameter uncertainty caused by structural breaks.

A large theoretical and empirical literature shows that ambiguity-averse investors respond to increased parameter uncertainty by reducing exposure to assets whose return distributions they perceive as unstable. Dow and Werlang (1992) for example, examine how uncertainty affects financial markets and discover that the behavior of investor changes in uncertain states as trading happens less frequently and there are larger no-trade intervals. Easley and O'Hara (2009) show that investors limit their stock-market participation when uncertainty about underlying parameters rises, while Boyle et al. (2012) find that investors tilt their portfolios toward familiar assets during periods of elevated uncertainty. These studies collectively suggest that uncertainty reduces willingness to hold risky assets and that investors substitute toward safer or more information-transparent securities.

Optimal portfolio-choice models that incorporate parameter uncertainty or at least account for estimation errors reinforce this intuition. Uppal and Wang (2003) show that allocations to risky assets decline when investors face ambiguity about return moments. Furthermore, Garlappi et al. (2007) demonstrate that the optimal portfolio of ambiguity-averse investors overweights safer assets and underweights the exposure to stocks with a high degree of parameter uncertainty relative to the standard tangency portfolio, which essentially "dilutes" the portfolio. Similarly, the results of Kan and Zhou (2007) imply that the optimal mean-variance portfolio is considerably less risky once the optimization acknowledges that means and covariances are subject to estimation error. Across these frameworks, uncertainty leads to systematic reductions in demand for assets whose parameters are imprecisely estimated, amplifying price pressure in precisely the situations where underlying parameters have changed.

This mechanism implies that periods of elevated parameter uncertainty should be associated with negative contemporaneous returns and higher expected returns going forward. When investors recognize that the return-generating process has shifted—but do not yet know the new parameters with precision—the collective reduction in demand depresses current prices. As compensation for bearing uncertainty is not paid instantaneously, expected returns rise until uncertainty gradually resolves.

Empirical evidence is in favor of this view. Stöckl (2020) for example, measures parameter uncertainty on the basis of a financial turbulence measure mentioned in Kritzman and Li (2010) as well as Stöckl and Hanke (2014) to predict the U.S. equity premium. By doing so, Stöckl (2020) shows that an increase in parameter uncertainty is a strong predictor of a substantial decline in aggregate U.S. stock market returns both in- and out-of-sample. In addition to that, Smith and Timmermann (2021) argue that break-aware models improve predictive accuracy out-of-sample while Smith and Timmermann (2022) introduce an instability factor based on the cross-sectional differences between stocks with a high and low exposure to time-varying risk premia.

The factor earns a premium that is comparable to well-known risk factors. Together, these findings indicate that parameter uncertainty is priced, both in the time-series and cross-section. Based on these observations, we argue that the literature supports the following mechanism: Structural breaks introduce uncertainty about the true parameters of the return-generating process. Ambiguity-averse investors reduce their exposure to assets whose parameters are uncertain. This collective response generates price pressure and negative short-horizon returns. As uncertainty resolves, expected returns remain elevated until investors regain confidence in the new regime.

This mechanism aligns directly with our empirical setting: if structural breaks induce parameter uncertainty at the level of individual stocks, then the time since the most recent break—break-age—naturally measures how much uncertainty remains unresolved. Young break-age corresponds to high parameter uncertainty and thus stronger price pressure, whereas older break-age reflects resolved uncertainty and a return of demand. This provides the conceptual link between structural breaks, ambiguity-driven investor behavior, and the cross-sectional pricing implications that our analysis investigates.

2.3   Parameter Uncertainty and Learning

Instead of viewing parameter uncertainty as a static risk factor, we leverage concepts from learning-based asset pricing models to motivate why break-age proxies for parameter uncertainty. The central insight taken from this strand of the literature is that investors do not know the true parameters governing returns but must infer them from realized data. In this framework, we consider structural breaks to represent "resets" of the learning process.

Early Bayesian frameworks as used in Veronesi (1999) show that investors update beliefs using noisy signals. Consequently, fluctuations in their confidence generate state-dependent volatility and expected returns Pástor (2000) formalizes this by showing that uncertainty about the mean return induces a learning premium that is highest when data is scarce and compresses as the sample size T grows. Pástor and Veronesi (2009) extend this to structural breaks in profitability, arguing that new regimes are characterized by high initial uncertainty and valuation ratios that are hypersensitive to new signals.

We argue that break-age serves as the empirical counterpart to the sample size T in these Bayesian frameworks. When a break occurs (i.e. break-age ≈ 0), the sample size that can be effectively used for estimation also collapses to 0. The observed data history thus becomes obsolete, and investors are forced to estimate parameters with a truncated sample. As break-age increases, the accumulation of post-break

data improves the precision of posteriors estimates and thus parameter uncertainty is resolving again.

Furthermore, recent general equilibrium work suggests this resolution is not instantaneous. Collin-Dufresne et al. (2016) show that learning about parameters like long-run growth rates generates endogenous volatility and long-run risk dynamics. When combined with ambiguity aversion (i.e. a situation in which investors distrust the data-generating process and entertain multiple priors) learning exhibits "information inertia". This inertia explains why price adjustments to structural breaks may be gradual rather than immediate, as investors require a threshold of evidence to confirm a regime shift has occurred (Leippold et al., 2008; Illeditsch, 2011). By capturing the time elapsed since the last regime shift, break-age directly proxies for the stage of this learning cycle, separating firms with "mature", well-understood parameters from those in the early, ambiguous stages of a new regime.

2.4 Measuring Stock-Level Parameter Uncertainty

Measuring parameter uncertainty on an individual stock level is notoriously difficult via traditional proxies, which often struggle with isolating the uncertainty from other, confounding signals. A standard and popular measures such as idiosyncratic volatility as introduced by Ang et al. (2006) as well as Ang et al. (2009), for example, suffers from conflating parameter uncertainty with risk or sentiment. Moreover, alternative measures such as the analyst dispersion as proposed by Diether et al. (2002) are subject to issues related to coverage and behavioral disagreement. Other proxies that have been used in the literature include synchronisity by Roll (1988) or beta estimation error as mentioned by Shanken (1992). While the former captures market friction in addition to parameter uncertainty, the latter is sensitive to short-sample instability.

In contrast to these existing proxies, break-age provides a conceptually distinct and economically grounded measure of single-stock parameter uncertainty that isolates uncertainty by capturing the obsolescence of historical information. Rather than inferring uncertainty from the width of the return distribution (risk) or the noisiness of the signal (frictions), break-age measures the length of the relevant sample path available for learning. It identifies the time elapsed since a discrete structural shift in the data-generating process, acting as a direct proxy for the data scarcity that drives uncertainty. By relying on statistically identified regime changes, break-age disentangles the quantity of information from the quality of the signal, offering a clean, history-dependent indicator of uncertainty that is unrelated to the trading frictions and behavioral biases plaguing existing measures.As such, it offers a clean alternative that isolates the uncertainty channel predicted by models of learning and ambigu-

ity aversion, and complements existing measures by capturing a form of parametric instability that cannot be inferred from returns or fundamentals alone.

3 Data & Methodology

In the following section we describe the dataset and methodology used to investigate how well break-age proxies for parameter uncertainty and whether or not there is a significant premium in the cross-section for stocks with recent breaks. Therefore, the first subsection briefly outlines the dataset underlying our examination. After that, subsection two describes the actual methodology we use in more detail.

3.1 Data

To evaluate how well break-age proxies for parameter uncertainty and how breaks are priced in the cross-section of stock returns, we use the following datasets: We focus on the US-stock market and use monthly delisting-adjusted stock returns from CRSP beginning in 1966 as this marks the year in which we start to have broad data coverage for a plethora of firm characteristics. Furthermore, we exclude stocks of companies with a share price below USD 5.0 as is also standard in the asset-pricing literature. Moreover, we also do not include the first 60 months of observations for stocks that had an IPO after 1966 to avoid diluting our results with the well-known return dynamics following an IPO. Based on these adjustments, our dataset still covers 19'828 individual PERMNO's. Furthermore, we download factor returns and risk-free rates from Kenneth French's website[1].

3.2 Methodology

This section details the empirical framework used to identify structural breaks and quantify their pricing implications. Thus, we first present and discuss the different breakpoint detection models that we use to identify structural breaks in stock returns. Thereafter, we explain how we calibrated these tests using extensive simulations to optimize the trade-off between detection power and false positives, ensuring robust identification in noisy financial data. Subsequently, we outline our event study design before we detail the construction of a rigorous counterfactual via Mahalanobis-distance matching which enables us to isolate the return premium associated with uncertainty.

---

[1]See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

3.2.1 Breakpoint Detection Methods

Before digging deeper into which breakpoint detection methods we use, we first want to clarify some terminology related to breakpoint detection in general. According to Lleo et al. (2020), breakpoint detection methods are defined as machine learning techniques that aim at identifying sudden changes in a given time-series $(Y_t)_{t=1,...}$ as represented by the dataset $y_{1:T} := \{y_1, ...y_T\}$. In this context, a breakpoint at time τ is identified "if the statistical properties of the sub-sequences $\{y_1, y_2, ..., y_\tau\}$ and $\{y_{\tau+1}, y_{\tau+2}, ..., y_T\}$ differ" (Lleo et al., 2020, p. 9). In the case of multiple breakpoints, a partition $\rho_{1:T}$ of the set $\{1, 2, ..., T\}$ needs to be identified such that the time-series' statistical properties within each of the sub-sequences remain the same while they change from one sub-sequence to another (Lleo et al., 2020).

Apart from defining what we mean with structural break and breakpoints in general, it is also necessary to further specify the amount of data we use for detecting breakpoints. In general, there are two possible methods regarding the timing of detecting breakpoints: First of all, there are so-called offline detection methods that use the entire available dataset $y_{1:T}$ to check for potential breakpoints. Alternatively, this type of method is also known as "batch change detection" or "phase I change detection". From a Finance perspective, such break detection methods are similar to in-sample test. Although we use a setting that is closer to out-of-sample tests for our actual investigation, a sound understanding of how offline methods work is essential for our paper. Thus, we further explore and illustrate the mechanics of offline detection methods using the following example: Assume that the exemplary time-series under investigation has at most one breakpoint. If there are no breakpoints, our time-series is independently and identically distributed according to F_0. In contrast to that, if our time-series does indeed have a breakpoint at any time τ, then the series follows the distribution F_0 prior to τ and thereafter follows the distribution F_1. Note that $F_0 \neq F_1$ (Ross, 2015). If we now want to test whether or not a break occurs at time τ, we can do so by testing the following hypotheses:

$$
\begin{aligned}
&H_0 : Y_i \sim F_0(y; \theta_0), \quad i = 1, ..., T \\
&H_1 : Y_i \sim \begin{cases} F_0(y; \theta_0) & i = 1, 2, ..., \tau \\ F_1(y; \theta_1) & i = \tau+1, \tau+2, ..., T \end{cases}
\end{aligned}
\tag{1}
$$

where θ_i refers to the parameters of the distribution that might be unknown. According to Ross (2015) such a problem can easily be solved with a two-sample hypothesis test. Of course, which test statistic should be used is determined by the assumptions that are made with respect to the distribution of the time-series and which parameter is expected to experience breaks. In the case of Gaussian distributed data and expected

breaks in the mean a suitable test statistic is given by the Student-t test. Once an appropriate test statistic $D_{\tau,T}$ has been chosen, we can determine its value and, given that $D_{\tau,T}$ exceeds a threshold value $h_{\tau,T}$, we can reject the null hypothesis of identical distributions and therefore conclude that a breakpoint has occurred at time τ (Ross, 2015).

In reality however, we cannot know in advance when a breakpoint occurs (i.e. τ is unknown to us). Hence, we have to compute $D_{\tau,T}$ for every point in time between $1 < \tau \leq T$ and chose the maximum value from there. This basically means that our time-series is split up into every possible combination of two sub-sequences and the two-sample test statistic is calculated for every possible splitting point. Following Ross (2015) the test statistic then becomes:

$$D_T = \max_{\tau=2,\dots T-1} D_{\tau,T} = \max_{\tau=2,\dots,T-1} \left| \frac{\tilde{D}_{\tau,T} - \mu_{\tilde{D}_{\tau,T}}}{\sigma_{\tilde{D}_{\tau,T}}} \right| \tag{2}$$

Consequently, we can reject the null hypothesis of no breakpoints if $D_T > h_T$ for an appropriately chosen threshold h_T that is determined by an acceptable probability for Type I errors α (i.e. the probability of falsely detecting a breakpoint although there is none). In terms of locating a breakpoint, the best estimate is given by τ, i.e. the value that maximizes D_T (Ross, 2015):

$$\hat{\tau} = \arg\max_{\tau} D_{\tau,T}. \tag{3}$$

The second type of detection method is referred to as online detection methods, "sequential change detection" or "phase II detection methods". In contrast to offline methods, online methods try to detect breakpoints at time τ, using only the data $y_{1:t}$ that is available up to t, where t is $0 < \tau \leq t \leq T$ (Lleo et al., 2020). Drawing parallels to Finance terminology, online detection methods correspond to out-of-sample tests. So whenever a novel observation y_t is added to the time-series, D_T is calculated again using the aforementioned offline methodology. In line with offline methods, a breakpoint is detected if $D_T > h_T$ which implies rejecting the null hypothesis of no breakpoints as long as the threshold value is chosen appropriately. If the threshold is not exceeded, another observation y_{T+1} is added and the procedure is repeated.

According to Ross (2015), the online setting allows the threshold value h_T to be chosen such that the probability of false positives (i.e. Type I error α) remains constant over time. Under the null hypothesis this implies that:

$$P(D_1 > h_1) = \alpha$$
$$P(D_T > h_T | D_{T-1} \leq h_{T-1}, \dots, D_1 \leq h_1) = \alpha, \quad T > 1 \tag{4}$$

11

In order to test whether structural breaks in the mean and/or variance can be detected, we use multiple different online breakpoint detection methods taken from the CPM framework of Hawkins et al. (2003) which provides standard (non-) parametric tests for single and multiple breakpoint detection. In total, we consider five separate parametric tests: three are based on the assumption of Gaussian distributed returns, and two are tests that do not require any assumptions regarding the distribution of returns. In detail, the tests we consider are the following:

- Gaussian:
  - Student-t test statistic for break in mean returns
  - Bartlett test statistic for break in variance of returns
  - Generalized likelihood ratio test statistic for breaks in mean and variance of returns

- Non-Gaussian:
  - Mann-Whitney test statistic for break in location of a series
  - Mood test statistic for break in scale of a series

All of these different tests are designed to detect single and multiple breakpoints and are implemented using the R package *cpm*. The basis for all our tests are simple stock returns in excess of the risk-free rate.

3.2.2 Hyperparameter Tuning

For all break-point detection tests, our goal is to balance false positive detections and the power to detect breaks at all. In order to optimally calibrate the breakpoint detection tests before investigating the data, we have thus evaluated the impact of several hyperparameters on the test performance over a large grid of simulated environments that span plausible features of return data.

In practice, we do not know the true data-generating process, the number or the exact timing of a breakpoint, nor the magnitude and direction of a break. To reflect this uncertainty, we have simulated over a grid of environments with changing features. These features include:

- Two stylized return processes: an i.i.d. Gaussian process and a weakly autocorrelated process
- Three possible regime-lengths: $n_{pre} \in \{60, 120, 240\}$ months

- Two regime structures: a single break at time $t = n_{pre}$ or a two-break regime in which the process changes from the baseline $t = n_{pre}$ and returns to the original regime at $t = 2n_{pre}$[2]

- Two types of breaks: a break in the conditional mean, where volatility remains unchanged and a variance break, where the conditional mean remains unchanged but the volatility changes.

- Six options for breaks in magnitude and direction: relative shifts of magnitude and direction: shift $\in \{-1, -0.5, -0.25, 0.25, 0.5, 1\}$[3]

Across all simulated environments, we evaluate each test under both a null scenario (no structural break) and an alternative scenario (with breaks of varying characteristics). In the null scenario, where the data-generating process remains constant throughout, any detected break constitutes a false positive. Thus, our primary metric is the Type I error rate, measured as the proportion of simulations in which the procedure incorrectly signals a break. In the alternative scenario, where the break timing, magnitude, type, and regime structure are varied according to the grid above, our main performance metric is the detection probability, i.e., the fraction of simulations in which a break is successfully identified. These two metrics jointly capture the key trade-off relevant for empirical applications: we are looking for a parameter setting that produces few false alarms in stable environments while maintaining strong detection power across a wide range of plausible data-generating processes and break configurations.

Table 1: Reports the average false positive rate (null scenario) and detection probability (alternative scenario) for each break-point detection method for the best hyperparameters for all tests.

Detection Method	Break-Type	Dist. Ass.	False Positives	Detection Probability
Student-t	Mean	Gaussian	15.33	51.08
Bartlett	Variance	Gaussian	3.00	36.05
GLR	Mean/Variance	Gaussian	6.83	59.36
Mann-Whitney	Mean	Non-Gaussian	13.33	43.74
Mood	Variance	Non-Gaussian	1.92	36.07

Across all methods, the GLR test provides the most favorable balance between false positives and detection probability, maintaining a relatively low false alarm rate (6.83%) while still detecting the majority of breaks (59.36%). The Student-t and Mann–Whitney tests offer higher detection probabilities but at the cost of elevated false positive rates, making them slightly less suitable under a conservative stance where spurious break detections would compromise inference. Conversely, Bartlett and Mood are highly

---

[2] Note that for a single-break regime, the total series length is $2n_{pre}$ and for a two-break regime it is $3n_{pre}$.

[3] For breaks in means, the post-break mean becomes $\mu_0(1 + \text{shift})$. The same is true for breaks in variances. Negative values correspond to a downward break.

conservative and produce almost no false alarms. This comes with markedly reduced detection probabilities. As a result, these test reliably avoid overstating structural change but are only likely to detect breaks when they are sufficiently large or abrupt. This underscores the inherently cautious nature of our calibration: by prioritizing robustness against false positives, we adopt settings that only flag structural breaks when the underlying shift is economically meaningful rather than marginal or transient.

3.2.3 Event Study

Our objective is to quantify how break-age (i.e the time since the most recently detected structural break) captures economically meaningful variation in parameter uncertainty and how this uncertainty is priced in the cross section of stock returns and over time. Structural breaks in the mean or variance of a return process undermine the reliability of pre-break data and temporarily amplify uncertainty about the true level of expected returns and risk. When such a break occurs, investors face elevated ambiguity about the new regime and should rationally require a higher discount rate or reduce exposure to the affected stock. As time since the break increases and additional post-break observations accumulate, uncertainty is gradually resolved, the precision of estimated moments improves, and parameter uncertainty declines. Break-age therefore provides a natural measure of the local amount of uncertainty associated with each stock at any point in time.

In this setting, treating each detected break as an event provides a flexible and conceptually consistent framework for studying how parameter uncertainty unfolds. Unlike traditional portfolio-sorting procedures, event-time aggregation does not require stocks to experience breaks at coordinated points in calendar time or in equal proportions. Because structural breaks are inherently heterogeneous and idiosyncratic across firms, the event-study approach allows us to track return dynamics cleanly around each break without forcing an artificial alignment into break-age deciles or long–short portfolios.

To characterize the market's reaction to breaks, we compute abnormal returns as the difference between each break-stock and that stock's matched benchmark portfolio. Specifically, for each detected break, we apply our matching algorithm to identify a set of comparable control firms based on a high-dimensional match in firm characteristics, performance history, and distributional similarity. For a detailed outline of the matching criteria and procedure, we refer to subsubsection 3.2.4. Abnormal returns for stock i at time t are defined as:

$$AR_{i,t} = R_{i,t} - R_{i,t}^{match} \tag{5}$$

14

where $R_{i,t}$ is the excess return of stock i at time t and $R_{i,t}^{match}$ is the equally-weighted return of its matched benchmark group. This construction avoids the need to estimate factor-model betas around the break which could turn out to be an operation prone to contamination when the return-generating process itself is shifting. Instead, our chosen approach ensures that expected returns are derived exclusively from economically similar firms experiencing no contemporaneous break.

For each break event, we then compute cumulative abnormal returns (CARs) by aggregating the matched abnormal returns across event time:

$$CAR_{i,\tau,\tau+k} = \sum_{s=\tau}^{\tau+k} AR_{i,s} \tag{6}$$

and obtain equally-weighted event-time CARs by averaging across all break events:

$$CAR_t = \frac{1}{N} \sum_{i=1}^{N} CAR_{i,t} \tag{7}$$

This design allows us to examine several conceptually distinct windows and helps us with disentangling different phases in the build-up and resolution of parameter uncertainty. Furthermore, the event-study framework directly connects break-age to observable return patterns and enables us to quantify how uncertainty shocks and their resolution are priced in the cross section of stocks.

3.2.4 Mahalanobis Distance Matching

To refine our benchmarking strategy for stocks experiencing structural breaks, we implement a rigorous matching procedure based on the Mahalanobis distance. This allows us to identify control stocks that are fundamentally similar in terms of observable characteristics but did not experience a structural break at the corresponding time. Using a Mahalanobis distance based matching procedure further aligns our methodology with recent, well-published matching-based event studies such as Daniel et al. (2023) and Daniel et al. (2024) to name just a few and thus provides a more precise estimate of the excess returns associated with break-induced uncertainty.

We implement the Mahalanobis distance based matching at the individual stock level with the objective of selecting the most comparable stock for each stock that experienced a break. The matching is updated on a monthly basis, conditional on data availability and a sufficiently long history without a break (i.e. minimum of 60 months), ensuring that matching occurs prior to the realization of the structural break. Formally, let $\mathbf{x}_i \in \mathbb{R}^K$ denote the vector of cross-sectionally normalized features of stock i, and let Σ represent the covariance matrix of the features over all eligible

candidates in the matching pool. We then define the Mahalanobis distance between each break-stock i and potential match j using the normalized form:

$$D_{i,j} = \sqrt{\frac{1}{K}(x_i - x_j)^T \Sigma^{-1}(x_i - x_j)} \qquad (8)$$

where K is the number of matching characteristics. This scaling factor ensures that the Mahalanobis distances are comparable across different sets of features and is similar in nature to the scaling used in Stöckl and Hanke (2014).

Each stock is matched on a set of pre-break firm-level characteristics that are commonly used in the cross-section of returns as in Nguyen and Phan (2017), Daniel et al. (2023) as well as Daniel et al. (2024). Specifically, we match on the log market capitalization, the book-to-market ratio, operating profitability, asset investment, momentum return over the prior 12 months and short-term reversal over the prior month. All variables are averaged over the past 60 months to mitigate noise and ensure stable estimates. Note that for our main results, we do not consider the industry as a matching criteria for two reasons: First of all, especially at the beginning of our investigation period, some industries only consist of a few companies that at times, experience breaks at the same time which further reduces the pool of viable candidates. Second, our results indicate that in many cases, a "better" match in terms of Mahalanobis distance is found outside the same industry.

Furthermore, we want to further iterate that the reason for why we explicitly match on pre-break characteristics rather than contemporaneous post-break characteristics is to avoid endogeneity.A structural break is, by definition, an event that may alter a firm's fundamental attributes (e.g., a break in variance changes the firm's volatility profile). If we were to match based on post-break characteristics, we would be comparing the treated firm to control firms that already resemble the new regime, effectively controlling away the very structural shift we aim to study. By matching on pre-break attributes, we ensure that we are comparing firms that were fundamentally similar ex-ante, allowing us to isolate the causal effect of the structural break itself.

4   Results

We present our empirical findings in three parts: First, we provide a comprehensive statistical characterization of structural breaks in individual stocks, documenting their frequency, magnitude, and detection delays. Second, we examine the market's reaction to these events, showing that the onset of a break triggers a period of negative abnormal returns consistent with a rise in ambiguity. Finally, we analyze the returns around the statistical detection of these breaks, providing evidence that the resolution

of uncertainty is accompanied by a sharp, positive revaluation, thereby confirming the pricing of the learning channel.

4.1 Breakpoint statistics

This section explores the statistics of breakpoints along three main dimensions. First of all, we investigate how common and frequent structural breaks in monthly stock returns are. Second, we investigate whether the direction of structural breaks is biased and how strong these breaks are on average. The third set of statistics is related to the detection and persistence of structural breaks.

Table 2: Presents statistics related to the prevalence and frequency of structural breaks in monthly stock returns across five detection methods. Each detection method name first indicates the type of break the method can identify. Mean refers to structural breaks in returns and Variance refers to breaks in volatility. Thereafter, the specific test is mentioned. The letters in parentheses denote distributional assumptions, where G refers to Gaussian and NG to non-Gaussian. The remaining columns show the relative number of stocks with at least one structural break, the total number of breaks for each detection method as well as the median and mean structural break per stock.

Detection Method	Stocks with Break (%)	Number of Breaks	SB per Stock Median	Mean
Mean - Student (G)	19.77	6096	1	2
Variance - Bartlett (G)	46.82	21925	2	2
Mean/Variance - GLR (G)	39.98	16517	2	2
Mean - Mann-Whitney (NG)	2.75	652	1	1
Variance - Mood (NG)	23.08	7516	1	2

Table 2 presents the prevalence and frequency of structural breaks in monthly stock returns across five detection methods. The names of the detection models are chosen to reflect the main properties of the test. Thus, each detection method name first indicates the type of break the method tries to identify. "Mean" refers to structural breaks in returns and "Variance" refers to breaks in volatility. Thereafter, the name indicates the particular test that is used to detect breaks. The letters in parentheses denote distributional assumptions, wher "G" refers to Gaussian and "NG" refers to non-Gaussian. Throughout the entire section 4 the same logic applies to the naming of the detection methods. The remaining columns show the relative number of stocks with at least one structural break, the total number of breaks for each detection method as well as the median and mean structural break per stock.

The results in Table 2 suggest that structural breaks in stock returns are not rare but widespread and economically relevant. Across detection methods, between 20% and 47% of stocks[4] experience at least one structural break in mean or variance if

---

[4]Note that after excluding stocks with a share price below 5.0 USD our entire sample consists of 19'828 individual companies that have enough history to be eligible for our investigation.

returns are assumed to follow a Gaussian distribution. For non-Gaussian detection methods, the prevalence of stocks is considerably lower ranging only from roughly 3% to 23%. While Gaussian tests aim at identifying changes in parameters, non-Gaussian tests pick-up on changes in the shape of or irregularities within a distribution, which is the reason why non-Gaussian tests show much lower prevalence. Put differently, the parametric tests detect systematic changes in moments. In contrast to that, the non-parametric tests focus on the entire shape of the distribution and are thus more conservative. For both types of tests, structural breaks in variances are more common than structural breaks in means. On average, monthly stock returns experience between 1 and 2 structural breaks in their unconditional mean or variance throughout the entire sample indicating that breaks are neither rare nor excessively common and potentially noise-driven events. In fact, the average number of structural breaks per stock implies that breaks reflect a persistent regime change which neatly leans into our notion of break-age capturing the parameter uncertainty caused by these infrequent but substantial changes in observable returns. Furthermore, we argue that neither the total number of breaks nor the number of structural breaks per stock would suggest that the detected breaks are simply caused by a misspecified data-generating process and the corresponding distributional assumption of the detection method.

Table 3: Presents statistics related to the direction and magnitude of structural breaks in monthly stock returns across five detection methods. Each detection method name first indicates the type of break the method can identify. Mean refers to structural breaks in returns and Variance refers to breaks in volatility. Thereafter, the specific test is mentioned. The letters in parentheses denote distributional assumptions, where G refers to Gaussian and NG to non-Gaussian. The table reports the proportion of structural breaks that are negative and positive together with the average post-break change conditional on the direction of the structural break. A break is considered to be negative when the post-break mean (for mean tests) or volatility (for variance tests) is lower than the pre-break value. Average changes are conditional on direction and test metric and are measured over the time window $t \pm 12$, excluding the break-month observation.

Detection Method	Break Direction (%)		Average Change	
	Negative	Positive	Negative	Positive
Mean - Student (G)	38.140	61.860	-0.137	0.178
Variance - Bartlett (G)	50.267	49.733	-0.161	0.190
Mean/Variance - GLR (G)	50.051	49.949	-0.168	0.201
Mean - Mann-Whitney (NG)	60.276	39.724	-0.143	0.166
Variance - Mood (NG)	51.051	48.949	-0.127	0.171

Table 3 presents statistics related to the direction and magnitude of structural breaks in monthly stock returns. More precisely, the table reports the proportion of structural breaks that are negative and positive together with the corresponding average post-break change conditional on the direction of the structural break. A break is considered to be negative when the post-break mean (only for mean tests) or volatility (only for

volatility tests) is lower than the pre-break value. Average changes are conditional on direction and test metric and are measured over the time window $t \pm 12$, excluding the break-month observation.

The results show that structural breaks in the means or variances of stock returns correspond to economically meaningful changes instead of statistical noise. The Gaussian tests related to changes in variance and mean-variance display a very balanced detection pattern of 50/50. This balanced ratio between positive and negative structural breaks indicates that there is no directional bias inherent to the testing procedure and thus supports the view that structural breaks mark genuine shifts in parameters. This notion is further backed by the fact that the conditional average changes are substantial in terms of magnitude. Average post-break changes in mean returns and variances compared to their pre-break counterparts range from 13% to 20% which qualifies as being economically meaningful and imply a significant re-pricing of these stocks. For the Gaussian tests, average negative changes range from -13% to -17%, while positive changes stretch from 17% to 20%. The variance test exhibits a similar adjustments but for volatility. gain these observations confirm that structural breaks represent material changes in the return generating process which in turn further reinforce a parameter uncertainty related interpretation. From the perspective of an investor, such abrupt changes are precisely the type of breaks that could make investors question their beliefs about their parameter estimates thus placing the investor in a state of elevated parameter uncertainty.

Table 4: Presents statistics related to the detection time and regime persistence of structural breaks in monthly stock returns across five detection methods. Each detection method name first indicates the type of break the method can identify. Mean refers to structural breaks in returns and Variance refers to breaks in volatility. Thereafter, the specific test is mentioned. The letters in parentheses denote distributional assumptions, where G refers to Gaussian and NG to non-Gaussian. The table reports the median and mean detection delay, defined as the number of months between the date to which a break is assigned and the date at which the method actually detects it. In addition, the table displays the median and mean duration between consecutive structural breaks for stocks experiencing at least two breaks.

Detection Method	Time to Detection		Time Between Breaks	
	Median	Mean	Median	Mean
Mean - Student (G)	2	8	3	38
Variance - Bartlett (G)	12	19	34	54
Mean/Variance - GLR (G)	14	22	35	57
Mean - Mann-Whitney (NG)	19	27	20	45
Variance - Mood (NG)	20	29	55	80

Turning the attention towards Table 4 we explore the statistics related to the detection time and regime persistence of structural breaks in monthly stock returns across five detection methods. The table reports the median and mean detection delay, defined

as the number of months between the date to which a break is assigned and the date at which the method actually detects it. In addition, the table displays the median and mean duration between consecutive structural breaks for stocks experiencing at least two breaks.

These results allow us to investigate the link between break-age, uncertainty and pricing. If breaks take time to be detected, investors are confronted with a period of elevated parameter uncertainty which could potentially give rise to a priced premium. As evident in the columns presenting the detections delay, it takes time to detect structural breaks. The Gaussian tests for variance and mean-variance (Bartlett and GLR) for example, have detection delays of 12 to 14 months, reflecting the fact that shifts in mean and variance require several months of new, post-break data to be detected on a statistically significant level. During this period, investors face elevated parameter uncertainty as neither the direction nor the magnitude of the new regime is known. It is exactly this time in which investors start to realize that a change in the pricing of a stock is about happen. Furthermore, the average time-span between breaks suggests that these regimes last for roughly 3 to 6 years in the case of structural breaks detected by a Gaussian method. Non-Gaussian tests seem to be slower in terms of detection time and discover substantially longer regimes. Taken together, the detection delay and regime persistence results presented in Table 4 provide further support for why break-age is a valid proxy for parameter uncertainty.

All in all, the results discussed in subsection 4.1 establish that the return-generating process of individual stocks is characterized by rare, sizable, and informationally opaque structural changes. From the perspective of investors, these changes cause uncertainty related to how the new parameters of the return-generating process look like. This uncertainty can only be resolved with the release of new observation which can only happen over time. This mechanism supports the role of break-age as a well-suited proxy for the level of parameter uncertainty investors face, and provides the empirical foundation for examining how uncertainty is priced in the cross-section of returns.

4.2 Measuring Uncertainty via Break-Age

After exploring various statistics related to the frequency, direction, magnitude and detection of structural breaks, we now focus on the return dynamics leading up to and following structural breaks under each detection method. For this purpose, Table 5 presents the mean cumulative abnormal returns in the months surrounding identified structural breaks in monthly stock returns across five detection methods. Abnormal returns are computed as return differences relative to an equally-weighted matching portfolio, constructed as described in subsubsection 3.2.4. τ denotes the month in

which the structural break occurs. The columns report mean cumulative abnormal returns over various event windows centered on τ alongside the their t-statistics.

The event-study results in Table 5 strongly suggest that structural breaks in return parameters are linked to pronounced short-term dynamics in abnormal returns that align closely with our breakpoint statistics. Across all Gaussian-type tests except the Student test for breaks in mean returns, we find the same consistent pattern in cumulative abnormal returns: (i) there are no economically meaningful cumulative abnormal returns in the 12 to 7 months preceding a structural break, (ii) there is a statistically and economically significant negative reaction at the time of the detected break, (iii) this negative performance deepens during the first 3 to 12 post-break months and (iv) gradually fades if longer horizons are considered. This development of cumulative abnormal returns perfectly matches our intuition related to the behavior of market participants and is exactly what we would expect if a structural break causes a noticeable increase in parameter uncertainty, followed by a period of gradual normalization as investors have the opportunity to reform their beliefs and process new information.

Table 5: presents mean cumulative abnormal returns in the months surrounding identified structural breaks in monthly stock returns across five detection methods. Each detection method name first indicates the type of break the method can identify. Mean refers to structural breaks in returns and Variance refers to breaks in volatility. Thereafter, the specific test is mentioned. The letters in parentheses denote distributional assumptions, where G refers to Gaussian and NG to non-Gaussian. Abnormal returns are computed as return differences relative to an equally-weighted matching portfolio, constructed as described in subsubsection 3.2.4. τ denotes the month in which the structural break occurs. The columns report mean cumulative abnormal returns over various event windows centered on τ alongside the their t-statistics.

Detection Method		$r_{\tau-12,\tau-7}$	$r_{\tau-6,\tau-1}$	r_τ	$r_{\tau,\tau+3}$	$r_{\tau,\tau+6}$	$r_{\tau,\tau+12}$	$r_{\tau,\tau+24}$	$r_{\tau,\tau+60}$	$r_{\tau+13,\tau+24}$	$r_{\tau+13,\tau+60}$
Mean - Student (G)	Mean	-0.002	0.095	0.005	0.027	0.077	0.151	0.159	0.269	0.045	0.201
	t-Stat	-0.143	3.215	0.250	0.499	1.171	1.810	2.094	1.445	1.211	2.430
Variance - Bartlett (G)	Mean	-0.005	0.024	-0.011	-0.017	-0.017	-0.009	-0.010	-0.004	-0.002	0.010
	t-Stat	-1.446	5.005	-3.617	-3.452	-2.769	-1.212	-0.865	-0.218	-0.189	0.574
Mean/Variance - GLR (G)	Mean	-0.001	0.044	-0.017	-0.032	-0.030	-0.023	-0.023	0.008	-0.008	0.027
	t-Stat	-0.188	6.329	-4.610	-4.721	-3.468	-2.031	-1.521	0.318	-0.870	1.350
Mean - Mann-Whitney (NG)	Mean	-0.106	0.236	-0.076	-0.198	-0.357	-0.480	0.209	-0.236	0.051	0.075
	t-Stat	-1.920	1.933	-2.038	-1.621	-1.973	-1.450	0.344	-0.312	0.306	0.423
Variance - Mood (NG)	Mean	-0.003	0.017	-0.012	-0.024	-0.022	-0.009	-0.026	-0.002	-0.023	0.006
	t-Stat	-0.589	2.756	-2.908	-2.959	-2.137	-0.654	-1.568	-0.089	-1.830	0.224

The absence of a systematic pre-break performance is an important because it suggests that the structural breaks are not caused by slow-moving trends or a deterioration of fundamentals over time. The fact that CARs remain economically small and mostly insignificant in the 12 to 7 months leading up to a break indicates that the market does not anticipate the structural break before the change actually becomes detectable in the data. In other words, the breaks are not linked to pronounced pre-event drift, which would undermine their interpretation as from an uncertainty perspective. Note that the statistically and economically significant positive cumulative abnormal re-

turns 6 to 1 month prior to the break-date do not undermine the uncertainty interpretation because they to do not imply anticipation. The detection methods we used do not detect the exact moment when the underlying process changes. They simply detect the moment when the sample has accumulated enough evidence to conclude that the process has changed in the past. So the break-date is the optimal point in the sample that splits two regimes from a statistical perspective. Statistically speaking, this refers to the point that splits the sample into two segments that minimize within-segment variance and maximize between-segment difference. The actual economic change can occur before the break-date. Put differently, the pre-break window from $\tau = -6$ to $\tau = -1$ is the transition period in which we would expect the regime to start manifesting but has yet to be confirmed.

At the break-date τ, statistically significantly negative abnormal returns can be observed for variance-based and joint-mean-variance tests with event-window t-statistics ranging from -3 to -5. This pattern corresponds well to the breakpoint statistics, where variance-increasing breaks are common and economically large, with average volatility increases between 17% and 20%. Such abrupt changes in dispersion are noticeable by investors which, now facing a more uncertain environment, demand immediate compensation, the absence of which causes downward price adjustments. In line with this perspective, the magnitude of negative CARs is largest for GLR. This is not surprising because the GLR-test also detects the largest and most economically meaningful parameter shifts in mean and variance jointly.

The short-window CARs following the structural break are further evidence backing our explanation. The CARs of the variance-based test for the windows $\tau, \tau+3$, $\tau, \tau+6$, $\tau, \tau+12$ are mostly statistically significantly negative with t-statistics often exceeding -3. These patterns mirror the structure of breaks uncovered in subsection 4.1: The breaks detected by Bartlett, GLR and Mood are not only frequent but large in terms of magnitude. Investors require time to process the implications of these parameter changes. The observation that the CARs become less pronounced and typically lose significance for the windows $\tau, \tau+24$ and $\tau, \tau+60$ is consistent with the notion that uncertainty spikes around the break-date and gradually decreases with the release of new information over time. This provides investors with the opportunity to process new information to better understand the new regime. Consequently, the time-dependent profile of post-break CARs displayed in Table 5 closely matches the temporal profile of parameter uncertainty as proxied by break-age.

There is one notable exception to this pattern: the results displayed for the Student test for structural breaks in mean returns. The CARs for this test are positive rather than negative in the short-run and there is no statistically significant return associated with the break-date. Fortunately, this deviation can be explained with the breakpoint

statistics presented in subsection 4.1. The Student test detects a disproportionately large share of positive mean breaks (61.90%). This implies that most detected structural breaks are associated with an increase in the unconditional mean return. Given that the underlying mean return increases, the pricing response is twofold: there is a positive drift component (potentially reflecting favorable news or fundamentals) and a component that is related to the associated uncertainty and typically negative. As these two components cannibalize one another, the resulting CARs are muted or just slightly positive on the break-date and subsequently. Moreover, the breakpoint statistics suggest that the Student test detects many small shifts in the mean that are too minor to generate strong market reactions. The fact that this method detects structural breaks for 19.8% of stocks implies that, in our sample , roughly 3'200 stocks[5] are affected by such breaks. If we combine this information with the number of breaks detected by Student, the average number of Student breaks per stock is approximately 1.56. An average detection delay of only 2 months and such a high frequency points towards the test being highly sensitive to small changes which explains the timid short-run post-break CARs we observe for this method. It is unlikely that large structural breaks in mean can be reliably detected with such short lags. A statistically significant detection of a large change would require more post-break observations.

In the longer run, the only significant post-break CARs can be observed for the Student test. Again, this is fully consistent with the breakpoint statistics. As already mentioned, Student identifies a disproportionate share of positive mean shifts and the average magnitude of these shifts is economically meaningful (+17.8%). Given the short median detection delay of 2 months, the post-break CAR windows capture a substantial part of the upward drift associated with these positive mean breaks. Thus, the positive long-horizon CARs for Student do not contradict the short-run uncertainty interpretation, but reflect that, for this test, the structural break itself frequently corresponds to an improvement in underlying expected returns. In additon to that, the Mood test exhibits a statistically significant negative CAR in the $\tau + 13, \tau + 24$ window, while earlier and later windows remain insignificant. This is not indicative of a second event but more likely reflects the long detection delay of rank-based tests (median 20 months) and the persistence of volatility regimes. When Mood detects a break, the variance shift is still unfolding, and the 13–24 month window captures a continued adjustment to elevated perceived risk.

Taken together, the CAR patterns reflect many results already discussed in the previous section. Tests that identify large and economically meaningful shifts (i.e. Bartlett and GLR) produce pronounced negative reactions that slowly fade as time passes. In contrast to that, test that detect smaller and predominantly positive shifts (Student)

---

[5]See footnote 4 for the total number of stocks in our sample.

yield muted or offsetting abnormal returns. Overall, these results are very supportive of our interpretation that structural breaks cause sharp increases in parameter uncertainty. On an aggregate level, the market reacts negatively to these spikes in uncertainty, triggering a price decrease as witnessed by negative post-break CARs that gradually vanish as break-age increases.

4.3 The Resolution of Uncertainty

To understand how investors respond to structural breaks in returns, we have a closer look at the abnormal returns around (i) the estimated break-date (τ), the interval between the break and its detection ($\tau + 1$ to $\tau^D - 1$), and (iii) the detection date (τ^D) which denotes the month in which the detection method has actually detected a structural break with sufficient statistical significance. The chosen split of dates is very informative because it matches the timeline of an uncertainty shock exactly: Structural breaks generate informational ambiguity immediately, whereas detection represents the formal resolution of that uncertainty. Table 6 summarizes the findings of this analysis for each detection method.

Starting with the break-date returns we observe immediate uncertainty effects. For 4 of the 5 detection methods, abnormal returns at the break-date τ are statistically significantly negative. These uniformly negative break-date CARs are consistent with an immediate market reaction to heightened parameter uncertainty. More precisely, when the return-generating process changes in mean or variance, investors face a temporary increase in uncertainty about true expected returns and risk. The observed magnitudes of the reaction varies across methods but is mostly likely due to their power and the types of structural breaks they detect. The two non-Gaussian tests (i.e. Mann-Whitney and Mood) yield more dispersed break magnitudes and lower break frequencies, generating larger point estimates but at the cost of lower precision. The only exception is the Student test, whose break-date abnormal return is economically small and statistically insignificant. For an explanation for why this is the case we refer to subsection 4.2 and the breakpoint characteristics presented in subsection 4.1.

From our perspective, the columns related to the cumulative abnormal returns between the structural break and its detection date $r_{\tau+1,\tau^D-1}$ is the most informative because we observe a remarkable pattern: For all Gaussian tests, CARs are insignificant or significantly negative. These results reveal that markets continue to price elevated uncertainty after the structural break has occurred but before it is detected. This implies that uncertainty does not simply or instantaneously resolve once the structural break happens but instead, investors gradually process new information to better understand the direction and magnitude of the regime change. Thus, this period is still associated with a return discount and increasing expected returns.

Table 6: presents (cumulative) abnormal returns around structural breaks and their subsequent detection across five detection methods. Each detection method name first indicates the type of break the method can identify. Mean refers to structural breaks in returns and Variance refers to breaks in volatility. Thereafter, the specific test is mentioned. The letters in parentheses denote distributional assumptions, where G refers to Gaussian and NG to non-Gaussian. The break-date τ is the endogenously estimated point at which the underlying return-generating process is assumed to have changed. The detection date τ^D is the first month at which the detection method identifies the break with statistical significance. The table reports abnormal returns at the break-date r_τ, cumulative abnormal returns between the break and detection dates $r_{\tau+1,\tau^D-1}$, and the abnormal return associated with the detection date r_τ^D. Abnormal returns are computed as return differences relative to an equally-weighted matching portfolio, constructed as described in subsubsection 3.2.4.

Detection Method	r_τ		$r_{\tau+1,\tau^D-1}$		r_τ^D	
	Mean	t-Stat	Mean	t-Stat	Mean	t-Stat
Mean - Student (G)	0.005	0.250	0.022	1.154	0.513	31.832
Variance - Bartlett (G)	-0.011	-3.617	-0.024	-4.250	0.191	32.045
Mean/Variance - GLR (G)	-0.017	-4.610	-0.022	-3.109	0.239	29.663
Mean - Mann-Whitney (NG)	-0.076	-2.038	-0.033	-0.360	0.107	3.701
Variance - Mood (NG)	-0.012	-2.908	0.071	5.308	0.084	15.321

Although the results for the non-Gaussian tests seem contradictory at first, they can be explained. The Mann-Whitney result is qualitatively similar to the Gaussian tests but noisier due to low power and a small number of detected breaks as documented in Table 2. The Mood test, by contrast, shows statistically significant positive CARs between break- and detection date, which most likely reflects its extremely long detection delays. Because Mood tests detect breaks on average more than 50 months after they occur, the initial uncertainty has most likely been resolved before the test detects the break statistically. As a consequence, the "between" CARs of Mood are not caused by the collectively ambiguity-averse behavior of market participants but rather reflect the normalization and stabilization period that follows the break, yielding positive CARs.

At detection date τ^D all detection methods exhibit large positive and highly significant abnormal returns. We argue that detection coincides with the moment when uncertainty resolves in the sense that the underlying parameter shifts become statistically clear. At this point, investors have gathered enough information to confidently estimate the direction and magnitude of the regime change to update expectations sharply and consistently, producing an immediate and substantial positive price reaction and thereby materializing the higher expected returns generated by parameter uncertainty. Given the fact that all five tests show large, positive and, highly significant detection-date CARs strongly supports the uncertainty-resolution perspective we have argued.

All in all, we argue that the patterns around the break-date, between break- and detection date, and the detection date itself, paint a coherent picture of how markets

price parameter uncertainty. Structural breaks trigger an immediate deterioration in abnormal returns, consistent with a sudden rise in uncertainty about the underlying return-generating process. This uncertainty remains elevated and is priced during the period between the break and its statistical detection, as reflected by the systematically negative CARs for all Gaussian tests with meaningful power. Once the break becomes detectable, uncertainty resolves sharply, and markets respond with large and uniformly positive detection-date returns across all methods. The fact that detection coincides with such strong positive abnormal performance, while the pre-break and between-break-and-detection windows exhibit the expected discount, supports our central argument: breaks generate parameter uncertainty, investors are averse to that uncertainty, and break-age captures exactly the timing structure of uncertainty spikes and their resolution. This sequence of price reactions, from uncertainty shock over uncertainty discounting by market participants to the resolution of uncertainty fits seamlessly into the broader narrative of parameter uncertainty and ambiguity-aversion discussed earlier. Furthermore, it validates break-age as an economically meaningful and empirically powerful proxy for investor-perceived uncertainty casued by structural breaks.

4.4 The Life-Cycle of Uncertainty

Table 7 further investigates the life-cycle dynamic of the uncertainty premium by reporting results for panel regressions that link the severity and timing of maximum drawdowns to the magnitude of structural breaks and the duration of post-break uncertainty. This allows us to dissect the anatomy of structural breaks, distinguishing between the immediate rational repricing of risk and the cumulative cost of parameter uncertainty. The dependent variables are the maximum drawdown between the break- and detection-date (Models 1–4), the time to maximum drawdown (Model 5), and the return from the drawdown to the detection-date (Model 6). ΔMean and ΔVolatility measure the change in mean returns and volatility between the 12 months before and after the break. Time to Detection denotes the number of months required for the detection method to identify the break. All regressions include stock and year fixed effects, and standard errors are double-clustered by stock and year.

First, columns (1) and (5) indicate that investors process fundamental risk efficiently. As shown in column (1), an increase in volatility (ΔVolatility) results in significantly deeper drawdowns. Crucially, column (5) reveals the speed of this adjustment. The massive negative coefficient for ΔVolatility of -22.84 indicates that volatility shocks drastically shorten the time to the maximum post-break drawdown. In other words, when the signal is strong (i.e. a sharp increase in volatility), the price crashes immediately to its new equilibrium.

26

Table 7: reports panel regressions linking the severity and timing of maximum drawdowns to the magnitude of structural breaks and the duration of post-break uncertainty. The dependent variables are the maximum drawdown between the break- and detection-date (Models 1–4), the time to maximum drawdown (Model 5), and the return from the drawdown to the detection-date (Model 6). ΔMean and ΔVolatility measure the change in mean returns and volatility between the 12 months before and after the break. Time to Detection denotes the number of months required for the detection method to identify the break. All regressions include stock and year fixed effects, and standard errors are double-clustered by stock and year.

	Max Drawdown				Time to Max Drawdown	Resolution Return
	(1)	(2)	(3)	(4)		
ΔMean	1.914***			1.871***	-11.429***	1.172***
	(12.029)			(11.933)	(-4.471)	(7.272)
ΔVolatility	-0.285***			-0.459***	-22.844***	0.930***
	(-4.770)			(-7.217)	(-7.899)	(10.643)
ΔMean xΔVolatility	0.272			0.192	16.262	6.368***
	(0.706)			(0.573)	(1.035)	(11.878)
Time to Detection		-0.005***	-0.011***	-0.009***		
		(-10.757)	(-12.865)	(-11.280)		
log(Market Cap)			-0.035***	-0.020***		
			(-6.809)	(-3.947)		
Time to Detection x log(Market Cap)			0.001***	0.001***		
			(6.145)	(4.182)		
Num.Obs.	14530	14530	14530	14530	14530	14530
Adj. R^2	0.181	0.092	0.107	0.225	0.161	0.280
Std.Errors	Stock & Year	Stock & Year	Stock & Year	Stock & Year	Stock & Year	Stock & Year
Test Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes

Second, columns (2) through (4) focus on the uncertainty premium itself. Even after controlling for the fundamental shift, the coefficient for "Time to Detection" (i.e. the break-age at the detection date) is negative and statistically highly significant across all models. This supports our notion of a "cumulative bleed": for every month the regime remains statistically ambiguous, the drawdown deepens by approximately 0.9% to 1.1%. In contrast to the instantaneous volatility shock, this repricing accumulates slowly and is consistent with investors demanding an increasing premium as historical data becomes progressively more obsolete.

Third, a closer look at column (4) reveals the mechanism driving this bleed. At a first glance, the negative coefficient for firm size suggests that large, efficiently-priced firms experience a more severe initial repricing, the positive interaction term between time to detection and firm size reveals that size acts as a buffer against the passage of time. Small firms, characterized by less liquidity, worse analyst coverage and information opacity, suffer from deeper drawdowns during the post-break and pre-detection phase. Contrary to that, large firms are effectively better protected from a prolonged repricing and thus offer a smaller uncertainty premium.

Finally, column (6) concludes the life-cycle of the uncertainty premium. The "Resolution Return" (i.e. the cumulative abnormal returns generated after the lowest maximum return until detection of the structural break is strongly associated with the magnitude of the initial volatility shock. The significantly positive coefficients of 0.93 indicates that the uncertainty premium accumulated during the drawdown is paid out as a sharp revaluation once the new pos-break regime becomes detectable.

5 Conclusion

In this paper, we propose break-age, defined as the time elapsed since the most recent structural break, as a novel and economically grounded proxy for single-stock parameter uncertainty. While the asset pricing literature has long recognized that investors face uncertainty about the true parameters governing returns, empirical progress has been hampered by the difficulty of isolating this uncertainty from fundamental risk, information frictions, or sentiment. We address this challenge by identifying the structural source of uncertainty: the sudden obsolescence of historical data.

Our analysis establishes that structural breaks in individual stocks are not rare statistical artifacts but frequent, economically significant events that fundamentally alter the return-generating process. By implementing a rigorous suite of parametric and non-parametric online detection methods that mimics the real-time information set of an investor, we demonstrate that these breaks induce a distinct life-cycle of parameter uncertainty. Specifically, we find that stocks experiencing structural breaks suffer sta-

tistically significant negative abnormal returns at the time of the break, followed by a period of continued negative drift while the regime shift remains statistically ambiguous. This discount is consistent with models of ambiguity aversion, where investors require a premium to hold assets with imprecise parameter estimates.

Crucially, we document that this uncertainty premium is distinct from standard risk compensation. The pricing dynamics reveal a clear "learning" channel: parameter uncertainty is highest immediately following a break as the effective sample size for estimation is small and resolves only when the break becomes statistically detectable. The sharp positive revaluation observed at the detection date confirms that the resolution of ambiguity is a priced event.

These findings have broad implications for empirical asset pricing and portfolio choice. By distinguishing between the width of the return distribution (risk) and the length of the relevant sample path (uncertainty), break-age offers a cleaner metric for testing learning-based theories than traditional proxies such as idiosyncratic volatility or analyst dispersion. Ultimately, our results suggest that the stability of the data-generating process is a key dimension of asset quality, and that the "age" of a firm's current regime is a fundamental determinant of its expected returns.

References

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns. *The Journal of Finance*, 61(1):259–299.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, 91(1):1–23.

Ang, A. and Timmermann, A. (2012). Regime Changes and Financial Markets. *Annual Review of Financial Economics*, 4(Volume 4, 2012):313–337.

Barberis, N. (2000). Investing for the Long Run When Returns Are Predictable. *The Journal of Finance*, 55(1):225–264.

Barsky, R. B. (1989). Why Don't the Prices of Stocks and Bonds Move Together? *The American Economic Review*, 79(5):1132–1145.

Bossaerts, P. and Hillion, P. (1999). Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn? *The Review of Financial Studies*, 12(2):405–428.

Boyle, P., Garlappi, L., Uppal, R., and Wang, T. (2012). Keynes Meets Markowitz: The Trade-Off Between Familiarity and Diversification. *Management Science*, 58(2):253–272.

Collin-Dufresne, P., Johannes, M., and Lochstoer, L. A. (2016). Parameter Learning in General Equilibrium: The Asset Pricing Implications. *The American Economic Review*, 106(3):664–698.

Cremers, K. J. M. (2002). Stock Return Predictability: A Bayesian Model Selection Perspective. *The Review of Financial Studies*, 15(4):1223–1249.

Dangl, T. and Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics*, 106(1):157–181.

Daniel, K., Klos, A., and Rottke, S. (2023). The Dynamics of Disagreement. *The Review of Financial Studies*, 36(6):2431–2467.

Daniel, K., Klos, A., and Rottke, S. (2024). Optimists, Pessimists, and Stock Prices. *Annual Review of Financial Economics*, 16(Volume 16, 2024):61–87.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22(5):1915–1953.

Diether, K. B., Malloy, C. J., and Scherbina, A. (2002). Differences of Opinion and the Cross Section of Stock Returns. *The Journal of Finance*, 57(5):2113–2141.

Dimson, E., Marsh, P., and Staunton, M. (2002). Triumph of the Optimists: 101 Years of Global Investment Returns.

Dow, J. and Werlang, S. R. d. C. (1992). Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio. *Econometrica*, 60(1):197–204.

Easley, D. and O'Hara, M. (2009). Ambiguity and Nonparticipation: The Role of Regulation. *The Review of Financial Studies*, 22(5):1817–1843.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643–669.

Garlappi, L., Uppal, R., and Wang, T. (2007). Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. *Review of Financial Studies*, 20(1):41–81.

Hawkins, D. M., Qiu, P., and Kang, C. W. (2003). The changepoint model for statistical process control. *Journal of Quality Technology*, 35(4):355–366.

Illeditsch, P. K. (2011). Ambiguous Information, Portfolio Inertia, and Excess Volatility. *The Journal of Finance*, 66(6):2213–2247.

Kan, R. and Zhou, G. (2007). Optimal Portfolio Choice with Parameter Uncertainty. *The Journal of Financial and Quantitative Analysis*, 42(3):621–656.

Kandel, S. and Stambaugh, R. F. (1996). On the Predictability of Stock Returns: An Asset-Allocation Perspective. *The Journal of Finance*, 51(2):385–424.

Kim, C.-J., Morley, J. C., and Nelson, C. R. (2005). The Structural Break in the Equity Premium. *Journal of Business & Economic Statistics*, 23(2):181–191.

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Hart, Schaffner and Marx, New York.

Kritzman, M. and Li, Y. (2010). Skulls, Financial Turbulence, and Risk Management. *Financial Analysts Journal*, 66(5):30–41.

Leippold, M., Trojani, F., and Vanini, P. (2008). Learning and Asset Prices Under Ambiguous Information. *Review of Financial Studies*, 21(6):2565–2597.

Lleo, S., Ziemba, W. T., and Li, J. (2020). Exploring Breaks in the Distribution of Stock Returns: Empirical Evidence from Apple Inc. SSRN Scholarly Paper ID 3700419, Social Science Research Network, Rochester, NY.

Markowitz, H. (1952). Portfolio Selection*. *The Journal of Finance*, 7(1):77–91.

Nguyen, N. H. and Phan, H. V. (2017). Policy Uncertainty and Mergers and Acquisitions. *The Journal of Financial and Quantitative Analysis*, 52(2):613–644.

Pástor, Ľ. (2000). Portfolio Selection and Asset Pricing Models. *The Journal of Finance*, 55(1):179–223.

Pástor, Ľ. and Stambaugh, R. F. (2001). The Equity Premium and Structural Breaks. *The Journal of Finance*, 56(4):1207–1239.

Pástor, Ľ. and Veronesi, P. (2009). Technological Revolutions and Stock Prices. *American Economic Review*, 99(4):1451–1483.

Paye, B. S. and Timmermann, A. (2006). Instability of return prediction models. *Journal of Empirical Finance*, 13(3):274–315.

Pesaran, M. H., Pettenuzzo, D., and Timmermann, A. (2006). Forecasting Time Series Subject to Multiple Structural Breaks. *The Review of Economic Studies*, 73(4):1057–1084.

Pettenuzzo, D. and Timmermann, A. (2011). Predictability of stock returns and asset allocation under structural breaks. *Journal of Econometrics*, 164(1):60–78.

Roll, R. (1988). R2. *The Journal of Finance*, 43(3):541–566.

Ross, G. (2015). Parametric and Nonparametric Sequential Change Detection in R: The cpm Package. *Journal of Statistical Software*, 66.

Shanken, J. (1992). On the Estimation of Beta-Pricing Models. *The Review of Financial Studies*, 5(1):1–33.

Smith, S. C. and Timmermann, A. (2021). Break Risk. *The Review of Financial Studies*, 34(4):2045–2100.

Smith, S. C. and Timmermann, A. (2022). Have risk premia vanished? *Journal of Financial Economics*, 145(2, Part B):553–576.

Stöckl, S. (2020). Parameter Uncertainty, Financial Turbulence and Aggregate Stock Returns. *SSRN Electronic Journal*.

Stöckl, S. and Hanke, M. (2014). Financial Applications of the Mahalanobis Distance. *Applied Economics and Finance*, 1(2):71–77.

Uppal, R. and Wang, T. (2003). Model Misspecification and Underdiversification. *The Journal of Finance*, 58(6):2465–2486.

Veronesi, P. (1999). Stock Market Overreactions to Bad News in Good Times: A Rational Expectations Equilibrium Model. *The Review of Financial Studies*, 12(5):975–1007.

Xia, Y. (2001). Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation. *The Journal of Finance*, 56(1):205–246.