Risks often change with little notice, rendering past observations obsolete. We can use dynamic mixtures of simple models to track change robustly. Still, tiny doubts about big outliers can make a huge impact on forecasts. We can’t fully predict this impact and will frequently disagree with each other on best approximations. That is why markets trade so much.

For all his genius, Albert Einstein never accepted the randomness inherent in quantum theory. He repeatedly denied that God would play dice with the universe (Born 2005). Modern physicists find these denials endearing, because it gives them an opportunity to feel cleverer than Einstein. Few doubt that chance is central.

Standard finance theory thinks it is clever because it allows for risk. But it tends to ignore the uncertainty enveloping risk, which is often the chance that most matters. Rarely does it note the dearth of relevant observations, much less the implications for pricing and risk management.

This chapter addresses an even deeper problem, namely that observations may cease to be relevant. In finance God doesn’t simply throw dice. Sometimes He changes His dice without telling us.

By dice I mean the regime that defines the relevant risks. With default risk the core die is a biased coin with chance $\theta$ of heads. Changing the die means changing $\theta$. We can introduce new parameters to describe the probabilities of regime change, and then let those probabilities change too.

Some regime changes are clear-cut. A coup brings to power a ruler who vows to repudiate the nation’s debt. Presto, credit spreads soar before a single new payment comes due. More likely there will be doubts. Sometimes
regime changes are purposely obscured. How many companies heading for bankruptcy advertise the fact beforehand?

Indeed regimes may change without emitting any immediate signal. The evidence accumulates only later, in outcomes that strain previous interpretations. These are the regime changes we will focus on here.

Unfittingly Fit

The counting game described in Chapter 6 has a fatal weakness. Neither $D$ nor $T$ ever decline. Hence, every time a particular $E = \frac{D}{T}$ frequency recurs, the variance $V \equiv \frac{E}{T}$ must have shrunk. In the limit, $V$ vanishes, making updates irrelevant.

As long as the underlying risk stays fixed, the observed frequency should eventually fit it extremely well. The tiny variance will reflect enough justified confidence that Bayesian statistics call the inverse of an estimator’s variance its precision. However, high precision is hard to reconcile with financial practice. Empirically, market behavior binds most $T$ to a few decades or less. And if the risk ever does materially change, we might need centuries to restore a close fit.

In other words, the model in the previous chapter doesn’t truly solve the problems it identified. It just defers them until $D$ and $T$ age. We need some potion to keep them young.

Could this problem be peculiar to beta or gamma distributions? Let’s check. Equation (6.2) relates $\Delta E$ to $V$ and unexpected news. The corresponding relation for $\Delta V$ works out to

$$\Delta V = \kappa_3 \frac{\text{news}}{\text{var}(\text{news})} - (\Delta E)^2,$$

(7.1)

where $\kappa_3$ denotes the third-order cumulant of beliefs, or skewness times the standard deviation cubed. Since $\text{news}$ by definition is expected to be zero, $V$ must be expected to decline. Indeed, if default risk stays constant, the following relation will hold:

$$\Delta \left( \text{precision} \left( \text{estimator} \right) \right) \equiv \text{precision} \left( \text{news} \right).$$

(7.2)

Hence, the problem isn’t the choice of distribution; it’s our Bayesian updating method. The precision of the estimator ideally grows with the pre-
cision of the latest news, regardless of the distribution or current conviction. It works great when default risk stays nearly constant and terribly when it doesn’t.

Evolutionary biology has long noted a similar dilemma. Darwin’s theory preached survival of the fittest. But if the environment changes, fit can become unfit. Fisher (1930) derived an equation analogous to (A6.2) (see the Appendix), showing that the mean genetic fitness should increase in proportion to genetic variance. He called it the fundamental theorem of natural selection.

High variance means that many genes, and the organisms expressing them, aren’t best suited to the current environment. Yet variance promotes more fitness in the long run. Optimal specialization depends on robustness, yet conflicts with it.

Mother Nature long ago devised an ingenious way to reconcile improvement with variation. It is known as sex. By mixing genes and keeping a recessive one in reserve, sexual reproduction maintains more variance than asexual reproduction. Indeed, if mutation and selection are slow enough the variance will stay approximately constant, a result known as the Hardy-Weinberg Principle after Godfrey Hardy (Hardy 1908) and Wilhelm Weinberg (Weinberg 1908).

Financial selection is far more destabilizing than sexual selection. But there must be some way to restore variance longer term. The only plausible explanation is mindfulness. Observers realize that past information may have lost relevance, and old inferences fade.

Updating can thus be viewed as a two-stage process. The first phase is standard Bayesian updating of beliefs given new evidence, assuming the risk regime $\theta$ stays the same. The second phase factors in beliefs about likely changes in $\theta$ and their impact.

Disturbingly Tiny Doubts

To illustrate the difference the second phase of updating can make, let us consider a situation we’ve all encountered: the suspicion that the game we’re playing isn’t fair. If outcomes revert to the mean we’ll feel reassured. What if they don’t?

To make this concrete, suppose we’re relatively confident a coin is fair, based on 500 heads in 1,000 observations, when new evidence arrives of 30 flips that all came up heads. If we model our uncertainty as a beta
distribution with parameters (500, 500), the new information raises our consensus estimate \( E \) to \( \frac{530}{1030} \approx 0.515 \). That is less than one standard deviation away from 0.5, which hardly seems to merit concern.

Or does it? Given a fair coin, the chance of no tails in 30 flips is one in a billion. Something doesn’t smell right.

As an alternative, let us posit a small \( \varepsilon \) chance, say one in a million, that the coin was switched before the latest sample. Suppose we’re also highly unsure what a switch might entail. To convey that, let us model the conditional risk as a beta distribution with fractional parameters, say (0.1, 0.1).

These tiny allowances suffice to raise \( E \) to 99.5% after 30 heads. Why the huge difference? Intuitively, 30 heads in a row make us highly confident a switch occurred to a heads-biased coin.

Hence tiny doubts can restore variance. However, they do so in a highly irregular way. If we tracked the evolution of beliefs in this example one toss at a time, our tiny \( \varepsilon \) would initially make little impact. After 15 heads in a row, \( E \) for the next round would be only 50 bps higher than if we ignored possible switches. But after ten more heads \( E \) would surge to nearly 0.95, or 4,300 bps above the baseline.

Naïve observers might think we had flipped out. Here are 25 heads, each with the same probability, but we get excited only about the last ten. We seem to have started out too complacent, scrambled to catch up, and possibly overreacted.

Strangely, our more suspicious neighbors, with an initial \( \varepsilon \) of one in a thousand, appear to have suffered similar mood swings. However, they reacted about ten heads sooner than we did. After 15 heads in a row they would lay 20 to 1 odds that the next round is heads, but they won’t need to because we think 2 to 1 odds are ample for betting tails. They will part us fools from our money, unless the fair coin really was fair and our caution turns out to have been justified. In that case our neighbors will look foolish for reading spurious patterns into randomness. See Figure 7.1.

This example is extreme. However, the qualitative features resonate throughout finance. They explain why financial views evolve much more quickly than nature does and with more spasms. Specifically,

- Doubts introduce new beliefs with high dispersion and new means.
- This increases the aggregate variance, which equals the mean variance of its components plus the variance of its means.
• If the regime has indeed changed, the realization will come slowly at first, quicken substantially, and then slow again near certainty.
• The combination can make the transition very steep, as in the curves in Figure 7.1. It is as if people ignore evidence until they hit a tipping point and abhor the uncertainty between.

Complexity Through Simplicity

Now we see how important it is to incorporate regime change into our models. But it is far from clear how. There are countless possible regimes, at least one for each default risk, and “countless squared” switching propensities between them. We can’t even write down the full set of equations, much less identify them, unless we impose major constraints. Since even small doubts about one parameter can have a big impact, our choice of constraints can easily deceive.

For example, suppose we identify two regimes, with $\lambda_{ij}$ the conditional switching rate from state $i$ to state $j$. The optimal updating equation for the conviction $p$ on the first regime is approximately
\[ \Delta p \equiv p(\theta_1 - E) \frac{\text{news}}{\text{var}(\text{news})} - p\lambda_{12} + (1 - p)\lambda_{21}. \] (7.3)

The first term provides the Bayesian updating. Since \( P(\theta_1 - E) = p(1-p) \) \((\theta_1 - \theta_2)\), it generates the S-curve effect of being slow at the edges and fast in the middle. The second and third terms give the regime-switching outflow and inflow, respectively.

The expected change will be zero at \( p^* \equiv \frac{\lambda_{21}}{\lambda_{12} + \lambda_{21}} \), where outflow balances inflow. If regimes switch rapidly enough, \( p \) will tend to stick closely to \( p^* \), in which case it’s hardly worth monitoring outcomes. Conversely, if switching is rare, \( p \) will usually be close to 0 or 1. These traits are shared by more complicated models. However, the strict band confining risk is more assumption than result, as is the possibility of describing risk evolution with a single variable. Moreover, equation (7.3) implicitly assumes that we know the switching rates, when in reality they’re very hard to identify precisely.

Returning to our simple counting game, let us figure out some simple adjustments for regime change. The easiest is to let past observations fade in relevance at a constant proportional rate \( \lambda \). That is, in every short interval \( dt \), old \( T \) shrinks by approximately \( \lambda t dt \) while old \( D \) shrinks by approximately \( \lambda D dt \).

If new observations enter at rate \( dt \), the net impact drives the relevant observation time \( T \) toward a constant value \( \frac{1}{\lambda} \), where rates of new evidence and fading relevance match. Near that value, each \( \lambda \) implies an estimator of the following form:

\[ dE \equiv \lambda \cdot \text{news} = \lambda(dx - Edt), \] (7.4)

where \( dx \) equals 1 with default and 0 otherwise. The long-run average \( E \) should match the average default rate, regardless of \( \lambda \). But \( E \) at any given time will hardly ever be right. Higher \( \lambda \) exacerbates the swings. Shortening the average information lag or effective observation period makes \( E \) less precise when risk is stable and more responsive when risk changes.

Another name for “estimator” is “filter,” because it screens information for presumed relevance. If the filter above runs infinitely long, it will form an exponentially weighted moving average rate of default. Let us call it an exponential moving average (EMA) filter for short. It is widely used.
in practical financial modeling; the preceding reasoning helps to justify it. Figure 7.2 depicts two sets of EMA weights.

The average information lag or duration for an EMA filter is \( \frac{1}{\lambda} \). Long-duration filters will behave quite differently from short-duration filters, and in practice we will rarely know which is best. So let us take as our aggregate filter a weighted average of different EMA filters and use Bayes’ Rule to update the weights. A small weight on EMAs of less than a few years’ duration corresponds to a dispersion of tiny doubts. An EMA with centuries-long duration implies much tighter confidence bands. Identifying an ultrasafe asset implicitly requires a duration of millennia.

To update the convictions \( p \) across different EMA filters, treat each filter \( i \) as if all beliefs were concentrated at its current mean \( E_i \), and apply equation (7.3) without any switching. That is, substitute \( E_i \) for \( \theta_i \) and set \( \lambda_{12} = \lambda_{21} = 0 \). Remarkably, dividing a complex aggregate update into combinations of simpler equations (7.3) and (7.4) does not distort the results. It is similar to how statistical mechanics divides the internal flux of various ensembles from the motion of centers of mass.

An example is charted in Figure 7.3. Four simple EMAs and a dynamic mixture respond to a stream of servicing interrupted by two defaults. For each simple EMA, the estimated \( E \) decays exponentially until default and
jumps at default in proportion to the decay rate. The dynamic mixture also alternates between continuous decays and jumps, but without fixed proportionality. If we rescaled the vertical axis in logarithms, the decay paths would be straight lines for simple EMA filters but curves for the dynamic mixtures.

Inferences from Markets

The aggregate filter behaves like an EMA with unsteady $\lambda$, whose values we can infer by fitting equation (7.4) to market data on $E$. When debt is being serviced, the effective $\lambda$ will change gradually, and usually downward as if a gamma-distributed belief were collecting new observations. On default the effective $\lambda$ will jump, and usually upward as if some past observations were suddenly rendered irrelevant. However, if we’re fairly confident default risk is high, $\lambda$ will start rising after extended servicing without default and jump down after default. Figure 7.4 provides a few examples.

More shapes are possible. The degree of confidence can vary in more ways than $E$ itself. It reflects centuries of history, yet also incorporates myriad beliefs about the future.
In all cases, the effective $\lambda$ is greater than the confidence-weighted average over individual $\lambda$, as the variance across different $E$ values also contributes. The discrepancy is even greater between the effective observation time (the inverse of the effective $\lambda$) and a confidence-weighted average of the observation times. These effects will make observers appear more myopic than they really are.

Dynamic mixtures of EMAs help justify the approximations used in the previous chapter. In the short term, beliefs behave as if they are gamma distributed with fractional $D$ and $T$ of a few dozen years or less. In the long term, dispersed beliefs about regime change offset the ratchet effects on $D$ and $T$. We can even account for default “discarding” more prior history than servicing does, without needing to posit emotional reactions to shocks.

Dynamic mixtures of EMAs also help explain the glue-like character of monetary expectations. Suppose we treat fiat money as debt that is repayable only through others’ voluntary rollover for material goods and services. Despite this dubious support, millions of exchanges every day testify to reliable servicing. Not only does $E$ drop for every individual EMA, but also confidence shifts toward shorter-duration filters. Within a few years, the aggregate $E$ can sink to seemingly infinitesimal levels. Yet if confidence is very short-term, evidence of monetary breakdown (or payments failure, if mediated through banks) might easily trigger panic.
Again this supports the notion that high-grade debt is inherently fragile. We’re highly confident in servicing, yet suspicious about our confidence. Neither borrowers nor regulators should take market confidence for granted.

Given the high uncertainty, it is tempting to make the market’s belief distribution our own. We can try to infer the structure from credit spreads at various maturities and options on those spreads, or simply by tracking current spreads very closely. We can then decide to believe those inferences, on the grounds that the market knows best.

No speculators fully believe this, as otherwise they will never find cause to trade. But they must believe it in part, both because the market’s view frequently triumphs over theirs and because it’s stressful to disagree. Hence nearly every speculator bends his views toward the market consensus. This shrinks the variance of beliefs and makes the consensus more stable, in line with equation (6.2).

Proponents call this deferring to the wisdom of the hive. That’s a bit misleading. The true wisdom of the hive comes from assimilating the foraging efforts of all its member bees. The variety and span of foraging is the hive’s variance. It presumes error, as few bees will be foraging in the best place. But variance is what the hive needs to adjust quickly to new information. If all bees cleave to consensus, the hive is doomed.

However, some cleaving to consensus is useful. Uncertainty can be debilitating. By helping stabilize prices, herding reassures that the future is manageable.

Unpredictability

The very complexity of the world saves us from forever believing alike. It makes it impossible for any observer to know exactly what the market believes now, much less how those beliefs will change tomorrow. Our predictors are bound to falter because of tiny doubts or errors that mount over time.

This follows mathematically from an extension of equations (6.2) and (7.1). Just as the change in mean is proportional to the variance, and the change in variance is proportional to the skewness, so too the change in skewness is proportional to the kurtosis. All these measures are known as cumulants. In general, each cumulant changes proportionally to the cumulant of next higher order, which gets progressively harder to identify and control.
I call this relationship Pandora’s Equation. It explains why market risk can never be tamed. Most puzzles in finance trace back to it.

The only remedy is to continually come back to the market, look for discrepancies between the behavior we predict and the behavior we observe, and readjust our beliefs about the market’s beliefs. In effect that’s what traders do when they study pricing charts. It’s not the useless exercise Samuelson (1965) and many other distinguished economists have claimed it to be. On the contrary, for short-term prediction it’s nearly always more important to study the market than to study economic fundamentals.

In that respect markets are like the weather. Whatever the forecast says, it’s a good idea to look outside and check. Almost surely the forecast will need some revision.

This isn’t just a question of timing, say, when a rainstorm will begin. Our best supercomputers can’t predict the weather more than a few weeks in advance, and even inside that horizon they occasionally make gross errors. Forecasters used to hope that better computers would muscle through. Lorenz (1963) discovered that miniscule errors in weather forecasting equations mushroomed into huge discrepancies. The discovery was seminal to what is known today as chaos theory.

Chaos theory shows that nonlinear deterministic equations can generate what seem like random results. Some invoke chaos theory, wrongly, to deny randomness in markets. However, our beliefs do have elements of chaos in that even perfect foresight of all future defaults won’t suffice to predict their evolution. They’re inherently turbulent.

This has huge implications for trading. Since each forecast bears the birthmarks of private beliefs and estimation errors, it frequently will disagree with other forecasts. That’s why financial markets trade so much.

Bear in mind too that expectations can affect real risks. If lenders think my default risk is high, I can’t borrow. But if I can’t borrow, I can’t readily change their minds. Conversely, if no one expects me to default, I can roll over my debts and defer default.

Turbulence makes financial markets harsh and unfair. Like the ocean or the winds, or the fickle Greek gods, financial markets can rage with tempest on little notice. When they do, they can wreck us or speed us to our destination. No wonder man fears them and tries to bring them under control.

However, we can’t confine financial markets to rigid boxes and shouldn’t try. Financial markets represent social imagination. They distill what we believe about the future, and what we believe others believe. We need to respect them and give them room to play.
Oddly, the world’s most famous critic of markets understood the power of imagination better than most market apostles:

A spider conducts operations that resemble those of a weaver, and a bee puts to shame many an architect in the construction of her cells. But what distinguishes the worst architect from the best of bees is this, that the architect raises his structure in imagination before he erects it in reality. (Marx 1887: chap. 7)

As usual, Marx overstated the case. In some ways that proves the point. No other species stretches imagination the way humans do, and then strives to make it real. That path can never be smooth.

“At last Osband is getting to the point,” said Prometheus. “Regime change under uncertainty is central to finance.”

“How does that distinguish finance from other fields?” asked Pandora. “Heraclitus observed 2,500 years ago that change was universal, and quantum physics has shown that all matter embeds uncertainty.”

“In finance the uncertainty is so big, and the relevant observations so few, that even the expectations can’t be measured precisely. Markets measure the beliefs about risk rather than risk itself. And risk changes often enough that beliefs rarely have time to converge on the truth.”

“Twentieth-century finance theory focused too much on risk, too little on changes in risk, and hardly at all on beliefs about changing risk.”

“Yes, that’s why it missed your equation. Each cumulant of beliefs gets updated according to the next higher-order cumulant. It makes for an endless ladder, except in the special case of perfect normality.”

“You give me too much credit, Prometheus. I barely understand what a cumulant hierarchy is.”

“You understand what it means—that financial markets will never completely predict their own evolution. That a lot of what we call forecasting is just dynamic tracking. That risk stays outside the box.”

“Will readers understand the connections?”

“I’m not sure. The cumulant hierarchy arises naturally from the Fokker-Planck equation. Physicists have known about it for generations. In fluid dynamics they call it the moment-closure problem, and consider it
essential to understanding turbulence. But I have never seen scientists apply it to beliefs.”

“You’re seeing it now.”

“He needs to elaborate. At the very least apply it to continuous random walks. There your equation takes an even neater form. The volatility of each cumulant of belief varies directly with the cumulant above.”

“Give him time, Prometheus. Let him motivate the questions better before charging off to answer.”