6

Safety in Numbers

Many credit markets behave as if they are inferring from only a few dozen years of relevant observations and fractional evidence of default. Their beliefs are modeled most plausibly as highly dispersed distributions, where the consensus might be wildly wrong. The uncertainty calls for much larger contingent reserves on top-rated credits than standard regulatory calculations recommend.

Modern bond markets are highly sensitive to perceived default risk. Credit spreads are normally measured in basis points (bps), which are one-hundredth of percentage points. While default might not trigger complete loss, the residual or salvage value of a defaulted claim is often less than half the nominal value. Ideally the market wants a 1 or 2 bps accuracy in its estimates of default risk per annum.

There is no way to achieve that degree of accuracy for the debtors we care most about. Knowledge is always rooted in directly relevant observation, and we don't have nearly enough. So analysts branch into indirectly relevant or possibly relevant observation and draw on their imaginations. They are lucky to obtain one-in-one-hundred part accuracy.

This problem is fundamental to financial risk analysis. Yet it receives scant attention. To redress the balance, let's start with a real-life example.

Real-Life Ignorance

The Soviet bloc imploded around 1990, give or take a few years. The countries that emerged from the wreckage were a motley crew. Some had strong

resources and human capital bases; some did not. Some had thriving market sectors before the collapse; some strangled markets down to the end. Some were committed to democracy and the rule of law; some stonewalled. Some had serviced foreign debts faithfully, some hadn't, and some hadn't borrowed at all.

Despite these differences, nearly all of the new authorities were desperate for cash. Their economies were plummeting, tax collection was in tatters, and needs for reconstruction were immense. Western governments and international agencies provided tens of billions of dollars in debt forgiveness and new lending. It wasn't enough, and much was squandered or stolen.

As the post-Soviet successor states found their footing, they reached out to private Western lenders. Many of the latter were amenable—indeed eager—toward winning new clients. But they did want to cover the default risks. How much were those?

Even big investment banks had to admit they did not know. They had few contacts in the area and were hesitant to project from other experiences. So they looked for new hires to assist them and clients. I was one.

For an economist, getting hired by an investment bank was a bit like Scarecrow receiving a diploma from the Wizard of Oz. However, Scarecrow rose to the occasion better than I did. I didn't know the relevant default risks. I didn't claim to know.

How I envied my colleagues focusing on Latin America. They tapped better data sets with more clearly relevant histories. They consulted with seasoned technocrats at treasuries and central banks. Analysts weighed in confidently not only on default but also on related risks like currency devaluation.

As the year unfolded, the markets grew increasingly anxious about Mexico's ability and will to maintain the peso's targeted exchange-rate band versus the dollar. However, most Wall Street analysts remained confident that the target band would hold. Sure enough, the band held—until it didn't. In just a few days in late December the peso lost one-third of its dollar value.

Only then did I realize that my fog of ignorance wasn't just mine. It permeates the market as a whole. That was my first taste of real market knowledge.

Default Uncertainty

Would that market data were as abundant and easy to categorize as subatomic particles. Every spoonful of matter offers a trillion trillion potentially relevant observations. Contrast that to a couple of dozen countries emerging from the Soviet bloc. Even if we understood precisely every economic, political, or ideological factor relevant to debt servicing, we would not know how they hang together. What objective basis does that give us for knowing default risk?

Imagine, for example, that we could find 100 years of debt-servicing experience exactly analogous to Slovakia's debt servicing in 1994 but independent of it. Suppose that three defaults occurred. The single most likely default risk is 300 bps. But wait, the bond market wants us to gauge basis points, and our experiment rules out fractional defaults. Hence it would seem fairer to infer a most likely default risk ranging from 250 to 350 bps. The uncertainty would span 100 bps. And that is just for a single year of default risk. The price of a Slovak bond maturing in ten years could wobble by several percentage points without making a fair-value analyst cry foul.

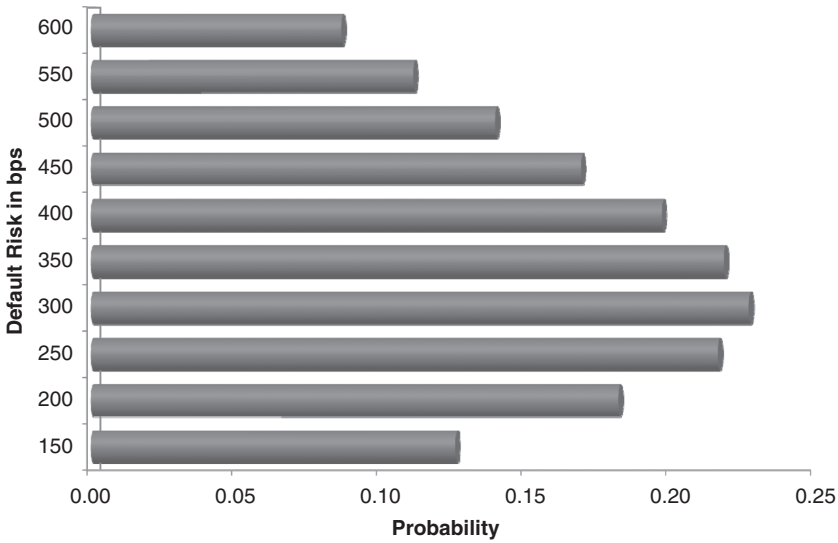

Figure 6.1
Default Risk Versus Probability of Three Defaults in a Century

Indeed, the uncertainty extends wider. A 450 bps default risk has a 20% chance of defaulting three times in a century. A 200 bps default risk has an 18% chance of defaulting three times in a century. So we shouldn't rule them out on the basis of this evidence alone. Indeed, any default risk between 150 and 540 bps is at least half as likely to trigger three defaults in a century as a 300 bps default risk is. See Figure 6.1.

Let us turn the question around. How many independent observations would we need to confidently identify a narrow range for default risk? The answer in most cases turns out to be hugely more than we have.

Suppose we want 99.9% assurance that our estimate lies within 10 bps of the underlying yearly default risk. If the true risk is 1%, we will need 100,000 independent yearly observations. If the true risk is 10%, we will need nearly ten times more. While credit cards and mortgages can generate those volumes of observations, only a fraction will share the same risk category and environment.

Is 99.9% confidence too stringent? Let's trim it to 95%. That demands barely one-third as many observations. Still the numbers are huge.

In sovereign debt, these thresholds are completely out of reach. If a few dozen countries experienced similar risks over a few dozen years, that yields only about a thousand yearly observations. A thousand yearly observations are unlikely to squeeze the confidence interval to less than 150 bps.

Hardly Any Defaults

Since we can't extinguish uncertainty, let's find ways to live with it. The simplest is to assemble T years of relevant independent observations, count how many D indicate default, and use the frequency D/T as an estimate of default risk. Since reasonable people will disagree over relevance and independence, both D and T will be fuzzy. But the market price will reflect a consensus weighted by wealth and conviction, so we can use that as a wobbly proxy.

Even without wobble, the estimate will move as new evidence comes in. Each step will look like this:

$$\frac{D}{T} \begin{array}{ll} \nearrow \dfrac{D+1}{T+1} & \text{if default,} \\ \searrow \dfrac{D}{T+1} & \text{if payment.} \end{array} \tag{6.1}$$

Let's suppose we gather *T* more years of evidence and it's all good: not a single indication of default. This will halve the estimated default risk. Hence, minimal halving time should provide a reasonable estimate of *T*.

In practice, it took only a few years for the best ex-Soviet performers to halve their initial credit spreads. If we generously accord each of them up to ten identical-but-independent comparators, *T* is at most a few dozen. In that case, the baseline *D* even for a relatively weak credit must be less than 3.

One might object that spread tightening reflected improvements in the economies and fiscal governance. This would make some reference defaults irrelevant. However, it would likely make many other observations irrelevant as well, shrinking *T*.

Even before the Soviet bloc collapsed, Hungary kept spreads tight despite many similarities to Poland, which defaulted. Evidently it persuaded markets that Polish default was barely relevant. Why then should subsequent Polish servicing be much more relevant? In general, while sovereign default often is contagious to spreads abroad, the impact usually fades within a year if others keep servicing. In effect, observations from other countries and periods get only fractional weight. Again, this shrinks the relevant *T*.

Even at fractional weight, enough related experience might trickle in to swell *T* overall. To estimate how much, let us consider what happens when a sub-100 bps credit risk suddenly defaults. It is hard to imagine its credit spreads not surging by 300 bps or more. Yet *D* is only 1 higher than before. Hence *T* must be only a few dozen, in which case the baseline *D* must have been much less than 1.

Bond markets' traditional emphasis on trust supports this interpretation. Good borrowers portray themselves as completely reliable, in hopes of issuing at minimal spreads. Any default is treated as disgrace. Yet if creditors completely believed borrowers, spreads would be zero. Only fractional evidence of default can reconcile trust with caution.

Consider too the difficulty in finding high-risk comparators. Leo Tolstoy wryly observed in *Anna Karenina* that while all happy families are alike, each unhappy family is unhappy in its own way. Trustworthy debtors are like happy families: they borrow cheaply, they repay promptly, both sides win, and the simple story repeats itself. Each defaulter, by contrast, has its own tale of why past failure should not be held against it now, as well as its uniqueness relative to others who have pledged the same and failed again.

Rational Updating

Fractional evidence of default is strange. Let's see if we can avoid it. Instead of positing a simple counting process for risk estimation, let us allow for varied beliefs and update them optimally given new evidence.

Bayes' Rule can accommodate a wide variety of behavior. At one extreme it can justify dogmatism. If I am 100% convinced default risk takes one particular value, applying Bayes' Rule won't change my mind. At the other extreme, it can justify a giant leap to conclusion. If I assign 50% probability of certain default and 50% probability of certain servicing, then the very first observation catapults me from maximum uncertainty to maximum conviction.

Still, Bayes' Rule implies some regularity in change. The most striking regularity concerns the mean default risk given beliefs. If we define *news* as the unexpected component of an observation, the following relationship always applies:

$$\Delta\,\mathrm{mean}(\textit{beliefs}) = \mathrm{var}(\textit{beliefs}) \cdot \frac{\textit{news}}{\mathrm{var}(\textit{news})}, \tag{6.2}$$

where Δ denotes the change induced by observation and var denotes variance. If we denote the mean belief by E and the variance by V, the preceding simplifies to

$$\Delta E = \begin{cases} +V/E & \text{if default,} \\ -\dfrac{V}{1-E} & \text{if payment.} \end{cases} \tag{6.3}$$

For example, if beliefs are roughly normal around the mean with a 50 bps standard deviation, then a century of reliable servicing will shave the mean by barely 25 bps, while three defaults in a row will boost the mean by less than 100 bps. Investors are a lot more skittish than that. Hence, market beliefs must be widely dispersed.

As long as E is small, the exceptional default will be far more newsworthy than servicing. The mean will jump when default occurs and slowly decay when it doesn't. The expected shift remains zero, which makes sense. If we observe what we expect, it shouldn't change our minds.

Initially we used equation (6.1) to model the updates on risk. Equation (6.3) will match this if $E = D/T$ and $V = \frac{E(1-E)}{T+1}$. Is there any probability distribution that has this property?

Indeed there is: a beta distribution with defining parameters D and $T - D$. The beta distribution is the most widely used probability model for variables spanning the continuum from 0 to 1. And its parameters don't need to be integers, just positive numbers. So our simple model with low T and fractional D looks more plausible than ever.

Figure 6.2 presents some charts of beta distributions with mean 1% and D ranging from 0.2 to 2. Even at $D = 2$, convictions are dispersed widely relative to the mean. Shrinking D increases the dispersion. For $D = 1$, convictions dampen exponentially as default risks rise. For $D < 1$, the mean is an uneasy compromise: most likely the true default risk is much lower or much higher.

As time between observations gets shorter, it is more convenient to work with the continuous-time counterparts known as gamma distributions. The relevant formulas are nearly the same. The Appendix provides details.

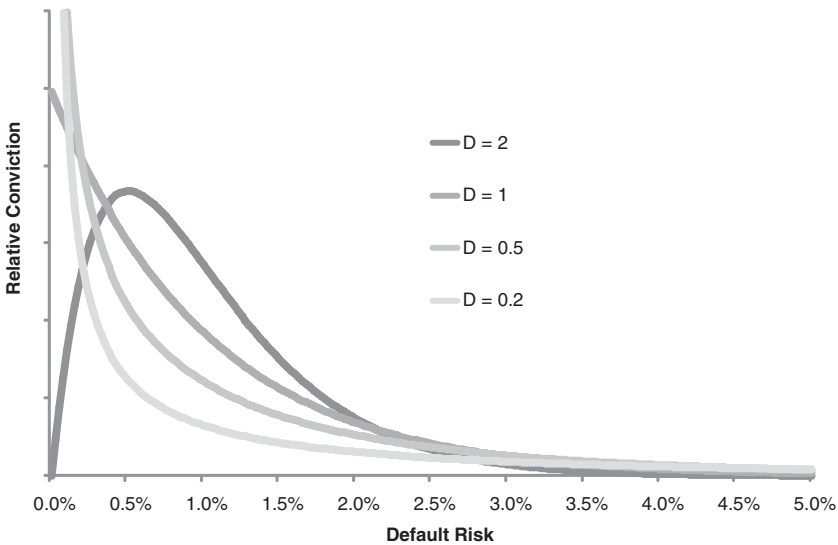

Figure 6.2
Beta Distributions with Mean of 1%

Contingent Reserves

Suppose a lender owns a basket of n credits considered identical in odds of default and independent in outcomes. Those qualities make them ideal for risk estimation. We pool them and similar credits to count D and T and form the same belief distribution for each. How should we estimate default risks for the portfolio as a whole?

The expected number of defaults is just n times the expected default risk E for each credit. Multiply by the average loss per default to estimate the expected losses out of nominal proceeds. Typically accountants will book a reserve to cover this, so as not to overstate lenders' anticipated profit.

Excess losses will chip away lenders' capital. To help stay out of trouble, prudent lenders will set aside an additional buffer, known as a contingent reserve. Ideally, the buffer should equate the marginal cost of excess reserves to the marginal cost of falling short.

If losses are distributed normally, a three standard deviation buffer provides 99.87% certainty of protection, while a four standard deviation buffer provides 99.997% certainty. For a small fixed default risk, the variance will roughly match the mean nE. Therefore contingency reserves are often judged adequate if they cover three or four times the square root of expected losses.

That judgment glosses over several problems:

- The risk environment might change, altering E. This is so important that the next chapter is devoted to regime change.
- Low D and low nE widen the standard deviation multiple needed for a given degree of protection. Intuitively, the fewer the expected number of defaults, the more scrunched most of the distribution will be against the origin and the fatter the tail risk of outliers more than three to five standard deviations from the mean.
- Lower D also fattens the tail, as defaults will be expected to cluster in high-risk scenarios. To take an extreme example, if we believe the joint default risk is either zero or 100%, then we expect to observe no defaults until everything defaults.
- Learning can change E substantially even if actual risks stay the same. This follows from equations (6.1) and (6.3). New evidence of default will boost E by a fraction $\frac{V}{E^2} = \frac{1-E}{D+E} \cong \frac{1}{D}$. For D small this will imply a huge percentage change in reserves.

Hence, a bank that leverages itself to the hilt can have more to fear from its low-risk assets than its high-risk ones. Low-risk assets reassessed at, say, 40 bps rather than 10 bps could double the standard reserves required for unexpected losses, and hence halve maximum leverage for a bank that relied on those assets to meet capital requirements. Longer-maturity assets are especially vulnerable, as they leave more time before redemption for observers to change their views.

This realization seems to have escaped the designers of the Basel Accords. Their differential capital requirements strongly favored nominally sound sovereign debt and mortgages. Lenders would naturally concentrate there, aggravating the credit cycle. Moreover, the regulations favored taking credit with the highest spread to nominal risk rating—the very credits the markets considered most overrated.

The extreme was the zero risk-rating for top-rated sovereign debt. In a way it's touching. Having lost the *droit de seigneur* and in some cases the *droit de seigniorage*, eminent sovereigns sought to assure the *droit de prêt*: the right to rollover debt indefinitely without a premium.

Many consider this a good idea, on the grounds that it builds trust and defers debt repayment to a richer future. They rarely stop to consider its costs. To begin with, most of these sovereign borrowers would deserve more respect if they pared down their debts. The last half-century has blessed most of these countries with peace, growth, and an enormous wealth and productivity advantage over the rest of the world. Their future looks comparatively tougher, even if peace prevails, because of demographic slowdown, resource depletion, and catch-up elsewhere. More structural fear would do them good, by encouraging them to save more and borrow less.

A second cost is abuse of some sovereigns by others. Many sovereigns, including Europe's PIIGS and several state governments in the United States, keep their top rating only because other sovereigns implicitly protect them. Helping the protected stuff their debt to banks doesn't reduce the protector's burden; it just obscures the connection. Eventually crisis forces the protector to choose between bank bailout and chaos.

A third cost is discouragement of bank lending to small- and medium-sized business enterprises and lower-rated large business. Their default rates are relatively high, so Basel rules require high contingent reserves. However, businesses are numerous and relatively diverse. This increases the relevant T in estimation and reduces the variance in aggregate performance.

In broad baskets of business loans, most unexpected risk comes from the credit cycle itself. Regulators should try to dampen this by encourag-

ing a smooth flow of credit to business. Pegging contingent reserves to recently observed default rates does the opposite.

Banks have a comparative advantage over bond markets in lending to smaller or weaker business, since they are better poised to get inside knowledge and better able to recoup costs of learning. Discouraging this in favor of sovereign lending or the origination of loans sold to third parties dampens real investment and reduces employment. While specialty finance companies step into the gap, their high returns suggest they tackle only the worst inefficiencies.

A fourth cost is concealment of risks through regulatory arbitrage. Regulatory arbitrage refers to repackaging that shrinks capital requirements without trimming actual risks. If capital requirements are set too high, regulatory arbitrage can benefit society. However, its very uniformity encourages concealment of the same kinds of risks in the same kinds of ways, until the very concentration poses huge systemic risks.

For example, under Basel I, a bank with a standard corporate loan would bear a capital charge of 8%. If it bought credit protection for this exposure from another bank via a traded derivative, the capital charge would sink to 1.6% (Servigny and Renault 2004). Basel II partially plugged that loophole by tying the capital charge to the insurer's credit rating. In doing so it opened a lucrative opportunity for top-rated insurance companies to offer credit guarantees that allowed banks to multiply their leverage.

The American International Group (AIG) was one of the few insurance companies that felt comfortable in this area, possessed a top rating, and faced little scrutiny from its external regulators or actuaries on derivatives exposure. So much banking system risk piled up at AIG that when it stumbled after Lehman's collapse, the U.S. government funneled $180 billion through it to bail out the banks AIG insured.

In fairness, the AIG episode points more to the dangers of high leverage and concealment than to the dangers of the particular forms that Basel II fostered. The main beneficiaries of AIG's bailout included some U.S. investment banks like Goldman Sachs that were subject neither to Basel II regulations nor the "Basel I plus leverage cap" applied to U.S. commercial banks. Also, some of the fallout could have been avoided by forcing more transparency into the credit derivatives market and requiring participants to post collateral for obligations likely due.

Rethinking Regulation

Financial risk regulation has created a kind of parallel universe. It aims to measure market risks better than the market does, with more objectivity and stability. It feels as though it works, except when it doesn't. When it doesn't, regulators invent new measures, which once again work until they don't.

Unlike Soviet-style planners, the masters of the parallel universe aren't antimarket. However, they share with Soviet-style planners an obtuseness to unintended incentives. Uniform rules for contingent reserve formation encourage uniform failings. Rigid risk pricing encourages imbalance and wasteful arbitrage. Decades of rewarding form over content have bred a culture of cynicism and irresponsibility.

This parallel universe desperately needs to be brought back to ours. The most important priority is to rethink the central aim of financial regulation. Is it to make everyone agree on risk? No, because that's impossible. Is it to help sovereigns and banks roll over their debt? No, unless one places them above the societies they are supposed to serve. Is it to decide what other lending is appropriate? No, because that's the job of lenders and credit markets.

Regulators should aim instead to dampen large fluctuations in aggregate credit. Market participants too easily confuse these with real changes in wealth and lose their bearings as a result. Regulators won't be perfectly clear themselves. But they are better placed than individuals to monitor credit markets and adjust policy accordingly. In spirit it's like adjusting money supply to stabilize inflation.

Some restraints on duration mismatch and leverage will be essential. Fixed caps are a lesser evil than no caps. In principle these caps ought to be related to aggregate credit growth, so that booms tighten leverage constraints while busts relax them.

Leverage caps would automatically assure some contingent reserves. As for the rest, lenders deserve more leeway to design their own policies. The basic mark-to-market test would be solvency, measured as an excess of assets over liabilities when valued at a moving average of market prices. Lenders having long-dated liabilities could average over much longer windows than lenders funding themselves via demand deposits.

The ultimate penalty for failing these tests would be "living death." Living death refers to prespecified procedures for addressing insolvency without the disruption or delays of ordinary bankruptcy. It would tap pri-

vate and public backup financing immediately and transfer ownership to the financiers. In some ways the U.S. government's Troubled Asset Relief Program did that by funneling credit quickly in return for equity stakes. It has been more effective and less costly than critics feared.

To supplement market valuations, and also to help markets appraise the lenders more reasonably, lenders would publish regular risk accounts. Risk accounts would look broadly like risk reports do now but with less pettifogging detail and with every important risk forecast accompanied by an estimate of its uncertainty. Risk accounting methodologies would require approval from professional associations and regulators.

One report that definitely has too much influence is perceived safety. Safety is usually highly uncertain, for statistical reasons regulators can't wish away. High uncertainty increases the reserve requirements for all assets but especially for the perceived safest. Tarashev (2009) makes a similar point.

Where the risk is well known, current regulations create perverse incentives to exploit it. Suppose a bank knows it's safe except for once in a blue moon, and that regulators don't fret blue moons. Then the bank is free to write a host of insurance contracts that capture premium every day and pay out in blue moons. Alternatively, if blue moons come too often for comfort, the bank can get back into regulatory grace with contracts that insure a fraction of failures but make the remainder far worse.

Some of the brightest minds in banking and structured finance are devoted to dubious estimation of unknown risks and intentional worsening of known risks. It is a huge disservice to society. When crises bring this into the open, the public is outraged and demands more regulation. Little does it realize how previous remedies worsened the problem. Better fewer regulations but wiser.

"Fractional evidence of default is odd," said Pandora. "But now that Osband explains it, I don't see any other way to reconcile empirical evidence with theory."

"I see it as a reduced form," said Prometheus. "If we take a mixture of beliefs, each with single default over different sample periods, the aggregate can resemble a beta distribution with fractional default."

"I prefer the simpler model. But I'll let him know; perhaps he can weave it into a later chapter. Do you have any other messages for him?"

"There's a new set of global banking regulations being drafted," said Prometheus. "Basel III. Ask him why he doesn't mention it."

"I already did. He said he doesn't know the details, and that key practical tweaks are still to come."

"They're introducing a leverage cap. They're raising capital requirements. They're demanding more backing for counterparty credit risk. They're demanding more liquidity coverage. Those are all good things."

"How about reducing procyclicality, improving risk models, and discouraging regulatory arbitrage?"

"They emphasize that as well. Only I'm not sure how. I do get a sense the Basel Committee is genuinely seeking new ideas. Plus there's some positive experience to draw on. Canada, for example, weathered the financial crisis a lot better than the United States or Europe did, thanks in part to better regulation."

"Do you think the reforms will turn the corner?"

"I'm not sure. It's going to be years before the new provisions are in place. A lot of big banks remain sorely overextended, particularly in Europe, and a PIIGS default would send them reeling."

"There will be huge political pressure to treat PIIGS debt as risk free."

"There already is. Likewise for the debt of California, New York, Illinois, and several other U.S. states that have borrowed well beyond their means."