

- MacDuffie, J. P. 1997. "The Road to 'Root Cause': Shop-Floor Problem-Solving at Three Auto Assembly Plants." *Management Science*: 479-502.
- MacDuffie, J. P., K. Sethuraman, and M. L. Fisher. 1996. "Product Variety and Manufacturing Performance: Evidence from the International Automotive Assembly Plant Study." *Management Science*: 350-369.
- Mehri, D. 2006. "The Darker Side of Lean: An Insider's Perspective on the Realities of the Toyota Production System." *Academy of Management Perspectives Archive* 20 (2): 21-42.
- Nakamura, M., S. Sakakibara, and R. G. Schroeder. 1996. "Japanese Manufacturing Methods at US Manufacturing Plants: Empirical Evidence." *Canadian Journal of Economics/Revue canadienne d'Economie* 29: 468-474.
- Ohno, T. 1982. "How the Toyota Production System Was Created." *Japanese Economic Studies* 10 (4): 83-101.
- Peters, T. J., and R. H. Waterman. 1982. *In Search of Excellence: Lessons from America's Best-Run Companies*. New York: Harper and Row.
- Shinohara, I. 1988. *NPS, New Production System: JIT Crossing Industry Boundaries*. Cambridge, MA: Productivity Press.
- Spear, S. J. 2004. "Learning to Lead at Toyota." *Harvard Business Review* 82 (5): 78-91.
- Spear, S. J. and H. K. Bowen. 1999. "Decoding the DNA of the Toyota Production System." *Harvard Business Review* 77(5): 96-108.
- Womack, J. P., D. T. Jones, and D. Roos. 1990. *The Machine that Changed the World*. New York: Rawson Associates.
- Zemell, E. 1992. "Yes, Virginia, There Really Is Total Quality Management." Anheuser-Bush Distinguished Lecture Series, SEI Center for Advanced Studies in Management, Wharton School, Philadelphia.

## 10 Assessing Risks

Thus far we have looked at risks as something that we want to anticipate—as in the chapters on indicators of disruption and distress—or to manage down—as in the Toyota production system's emphasis on reducing variability. Some investors, such as Berkshire Hathaway, claim, as their risk management strategy, that they only buy into low-risk businesses. Warren Buffett, Berkshire's chairman, said in his 1996 letter to shareholders: "I should emphasize that, as citizens, Charlie [Munger, Berkshire's vice chairman] and I welcome change: Fresh ideas, new products, innovative processes and the like cause our country's standard of living to rise, and that's clearly good. As investors, however, our reaction to a fermenting industry is much like our attitude toward space exploration: We applaud the endeavor but prefer to skip the ride."

But how do we know how much risk a business has? Traditional finance theory has one answer: in a world of efficient markets, a business's risk is captured in its beta. Beta—reflecting *systematic risk*—is the covariance of the business with the market<sup>1</sup> divided by the variance of the market. That part of a business's variance that does not covary with the market—called *specific risk*—can be diversified away by investors. But there is a growing literature, even in finance, suggesting that specific risk does matter. Therefore, another measure of risk is to consider both systematic risk and specific risk, or *total risk*.

From an operations perspective, there are two types of total risk. The first is process risk, which is recurrent; the second is better viewed as one-off. An example of process risk is the risk of machines breaking down. This can affect the value of a business through higher working capital (inventory) and reduced productivity. An example of one-off risk is the risk that a patent application will not be approved. This affects the value of a business directly.

These definitions raise the question of how we can measure the different types of risks, which we address in this chapter:

- **Systematic risk.** A convenient way to measure this is to return to the definition of beta as the covariance between the expected return from our business and that from the market portfolio. We assess risk mostly in order to value a business. It is vital to directly look to

the market for the value of a similar business, rather than try to calculate systematic risk without market validation. Then we expand on our calculation to see how operational leverage increases systematic risk.

• **Total risk.** We describe how you can use a risk map to visualize the severity and probability of changing risks in a business.

**One-off risk.** We describe the value-at-risk (VaR) framework to assess this type of risk. The VaR spells out the amount that we might lose (severity), for a given probability and time horizon.

**Process risk.** We describe how we can characterize the variability of this type of risk, with coefficients of variation, and how variability reduces productivity and therefore the value of the business.

This typology of risks helps clarify differences that confound even seasoned investors. For example, Warren Buffett declares that “we define risk, using dictionary terms as the ‘possibility of loss or injury.’ Academics . . . like to define risk differently. . . . In their hunger for a single statistic to measure risk, however, they forget a fundamental principle: it is better to be approximately right than precisely wrong” (Berkshire Hathaway 1993 Annual Report). But it is possible that Buffett is referring to one-off risks, while in finance theory risk is viewed more like process risk, in the ongoing ups and downs of the value of a security.

Furthermore, the typology reveals similarities and differences in the treatment of risks in finance and in this book. For example, finance theory suggests that risks from many sources in a portfolio can be diversified away. In operations, we too model how risk pooling can be helpful—e.g., in reducing inventory needed to serve a given level of demand. But there are also risks, such as process risks, in which the risk sources are in various stations along a process. In this case, the risks cascade and get amplified through the process. We should verify how the amplification could seriously affect our assessment of the risks.

### 10.1 Why Is Risk Assessment Difficult?

Before we get into some tools for assessing risks, it is helpful to ask why it is difficult to assess risks in the first place.

Suppose we were to ask the likelihood of facing a shortage in a critical component for production. One way to do this is to simply count the number of past shortages and divide by the time covered (say, 54 shortages in the last 3 years, or 18 shortages per year). This sort of risk assessment has several problems.

First, it does not account for severity. Some shortages are short, others longer. A histogram of shortages by severity would be useful. Second, even the histogram might convey a false impression of the distribution by severity. The histogram represents data from a

sample, but the actual distribution from the population of shortages might be very different—e.g., skewed toward severe shortages. Third, severity also needs to account for our ability to manage the shortage. A production line that can respond effectively to a shortage will find it less severe than another that is crippled by the same shortage. Fourth, a probability measure does not account for seasonality within a year, or a month or week. For example, the component may be more likely to run short at year end than at other times. Fifth, a probability measure also does not account for trends. Suppliers’ reliability changes over time. The number of suppliers changes over time. Our responsiveness changes over time. Sixth, the component probability does not consider how failures elsewhere would affect this probability, and how shortage in this component might impact failures elsewhere.

What this discussion suggests is that, given the complexity of the origins and implications of risks, any single measure will have its limitations. A proper risk assessment needs to account for the cost and benefit of undertaking the assessments, and the risk of not doing so.

### 10.2 Systematic Risk

One way to measure systematic risk of a business is to go back to the definition of beta:

$$\text{beta} = \text{Cov}(r_{\text{business}}, r_{\text{market}}) / \text{Var}(r_{\text{market}}),$$

where  $r_{\text{business}}$  is the expected return from the business in question and  $r_{\text{market}}$  that from a chosen market portfolio. Although the definition asks for expected returns, a reasonable starting point is to approximate these using historical returns.

Once we have determined beta, we can calculate the discount rate we will use to value the cash flows of our business:

$$\text{discount rate} = \text{risk-free rate} + \text{beta} \times \text{risk premium},$$

where the risk-free rate might be the treasury note return and the risk premium is the market return less the risk-free rate. Intuitively, the risk-free rate captures the time value of money, and the (beta  $\times$  risk premium) term captures the riskiness of cash flows.

But in reality, things can get messier. For one, there is “Roll’s critique” (named for Richard Roll of the University of Chicago): beta is very sensitive to which market portfolio we have chosen. So it seems wise to construct beta from a number of candidates, such as the Dow, S&P 500, NYSE Composite, Wilshire 5000, or the Morgan Stanley Capital Index.

Second, the discount rate may change over time. In their authoritative textbook, Brealey, Myers, and Allen suggest a way to address this: assume that beta  $\times$  risk premium is very slow-changing, so that we account for the change in discount rate by tracking the change only in the risk-free rate.

Third, we bear in mind that the market-derived values are usually more accurate than values derived bottom-up using covariances and variances. This is because the former use a more diverse set of information. So before we start calculating covariances and variances to find beta, we should see if there is a peer business with similar risks. "Similar risks" are of course hard to define, and we may only be postponing our problem. One way is to identify companies in the same industry with the same operating characteristics (using factors such as inventory turn or days of accounts receivable, as discussed in section 1.3.1). We can then check our bottom-up derivation against the published beta of our similar companies.

But how does systematic risk, or beta, relate to operations? It turns out that operational leverage—the proportion of assets that is fixed—increases systematic risk. We first give the formulaic explanation for this, then the intuitive sense of it.

The present value (*PV*) of an asset may be written as follows, since present values are additive (see Brealey and Myers or any standard finance text) and the cash flows from an asset is just revenue minus the two types of costs:

$$PV_{asset} = PV_{revenue} - PV_{fixed\ cost} - PV_{variable\ cost}$$

Therefore, the beta of an asset can be decomposed into betas of revenue, fixed cost, and variable cost:

$$\begin{aligned} \beta_{asset} &= \frac{PV_{revenue}}{PV_{asset}} \beta_{revenue} - \frac{PV_{fixed\ cost}}{PV_{asset}} \beta_{fixed\ cost} \\ &\quad - \frac{PV_{variable\ cost}}{PV_{asset}} \beta_{variable\ cost} \end{aligned}$$

Since fixed costs are fixed (by definition),  $\beta_{fixed\ cost}$  is zero. Also,  $\beta_{variable\ cost}$  is about the same as  $\beta_{revenue}$  since both are covariances with market returns, up to a fixed-cost constant. The above simplifies to:

$$\beta_{asset} = \frac{PV_{revenue} - PV_{variable\ cost}}{PV_{asset}} \beta_{revenue}$$

Again, using the identity that  $PV_{asset}$  is the present value of the revenues minus the costs, we can rewrite the above as:

$$\beta_{asset} = \left[ 1 + \frac{PV_{fixed\ cost}}{PV_{asset}} \right] \beta_{revenue}$$

The term  $PV_{fixed\ cost}/PV_{asset}$  is a measure of operational leverage, and the above shows that  $\beta_{asset}$  will increase as leverage increases, given  $\beta_{revenue}$ .

The intuition behind this is that with a higher proportion of assets that are fixed, a business is less able to cope with the vagaries of demand. When market demand is low, the

business incurs fixed costs, so its cash flows vary more with the market than they would for a business with a lower proportion of fixed costs.

As mentioned earlier, systematic risks are risks that cannot be diversified away. In the discussion above, even firm-specific factors that change the ratio of fixed costs to assets—perhaps originating from managerial competence or technological advances—would affect  $\beta_{asset}$ . But it may be hard to see other risks—say from information systems that might malfunction or human capital that may defect to the competition—that might also influence operational leverage. In other words, if we stick to just measuring systematic risk, theory does not give us a good indication of which risk is systematic and which is specific. Further, a growing amount of research suggests that even firm-specific diversifiable risks may be important in valuing companies. Finally, if we take the perspective of a manager, and not an investor who can diversify away specific risks, then total risk matters.

Taken together, the above suggests that we would do better to consider both systematic risk and specific risk, or total risk.

### 10.3 Total Risk

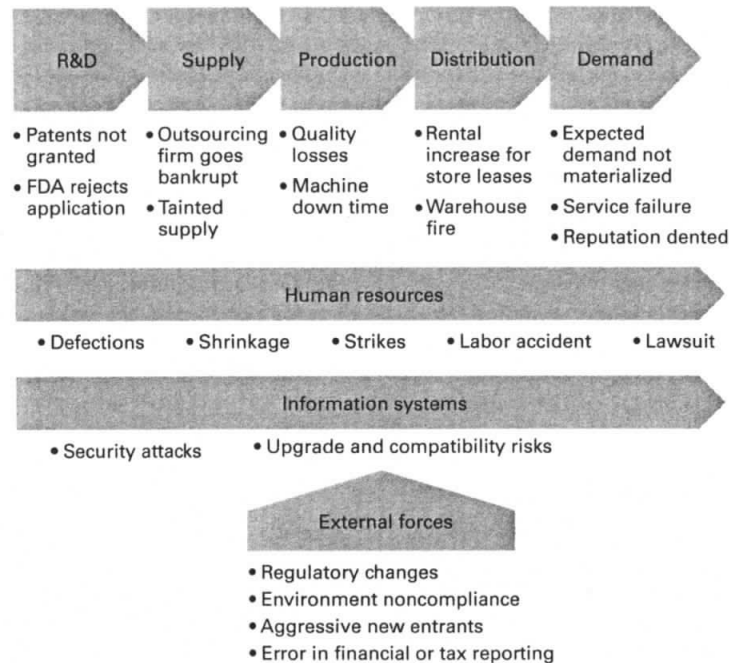
Not surprisingly, there are many ways to characterize the total risk of a business. As before, one useful approach is to appeal to the market. An especially helpful proxy is the divergence among analyst estimates of the business's earnings or sales. Of course, this measure is subject to some limitations—for example, it would not be useful if the business is covered by very few analysts.

At the opposite extreme, we can build a theoretical list of the risks that may arise. One useful framework is to identify risks that may interrupt the smooth addition of value along the value chain—from purchasing and R&D through production to distribution, as well as supporting functions such as human resource and information systems management. Figure 10.1 shows an example, and figure 10.2 shows an even more detailed list of risks in the area of production alone, from Shell's enormous "Comprehensive List of Causes" (ignore the details, which are not visible; the point is that Shell takes risks comprehensively).

We next need a tool to make these lists manageable. One useful tool is the "risk map" that prioritizes the risks for more detailed examination. An important variant is "failure mode and effects analysis" (FMEA), developed by the US military after World War II. In applying this to commercial situations, we focus on two dimensions of risks: the probability of occurrence and the severity of consequences.

#### 10.3.1 Probability

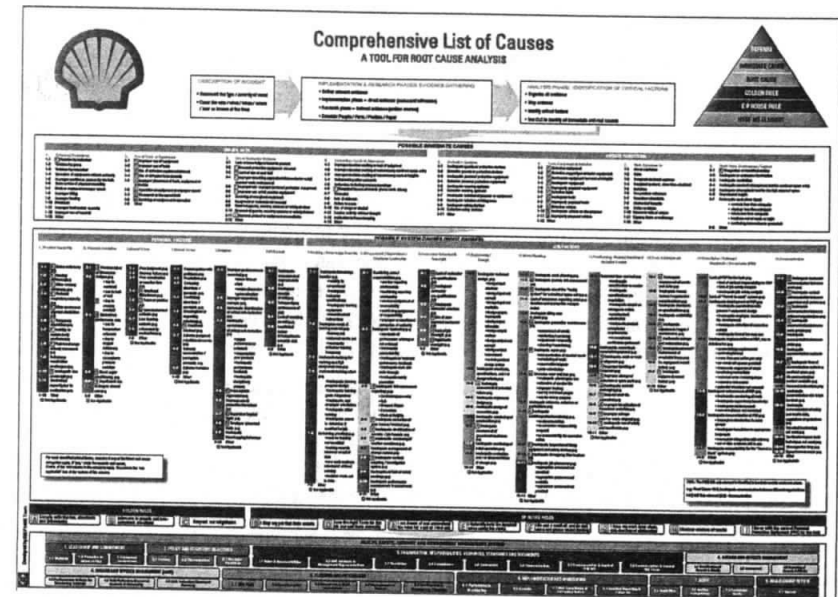
The probability of a risk is often given as a number, but that should be construed as the mean from a more complete characterization of probability, in the form of a probability



**Figure 10.1**  
Risks in a value chain.

distribution. The confidence of our probability judgment can be measured as the variance of the distribution: the smaller the variance, the more confident we are. This variance depends on how much information we can employ—and the cost of getting that information. For example, Sport Obermeyer, a skiwear maker, uses a six-member committee of purchasers, salespeople, and top managers to make probability judgments about the demand for new skiwear. Empirically, it has been found that this produces a superior demand forecast—in terms of both the bias (the forecast error is on average about zero) and confidence (the forecast error is proportional to the variance of the members' forecast).

In finance, probabilities (or variances) may be reduced through diversification. That is also true in operations—e.g., in risk pooling of workstations. But in operations, probabilities are not only correlated but also amplified. A workstation with a probability of downtime can amplify the probability of downtime at the next workstation. The probability that



**Figure 10.2**  
Shell's "Comprehensive List of Causes." (Details are not meant to be visible.) © Shell.

a supplier sends tainted supplies and the probability that that supplier goes into bankruptcy may be very high. Indeed, one may cause the other in a vicious amplifying cycle. We address these sorts of relationships between risks below.

### 10.3.2 Severity

The severity of a risk also depends on three factors. First, of course, there is the raw severity: how bad or what loss amount would be incurred if the risk materializes. Second, this severity depends on the probability and timing of detection; a high probability of early detection could mitigate the raw severity of the risk. Third, the ability of the organization to respond to the risk also mitigates the raw severity. For example, a flexible organization with some redundancy can shift resources to accommodate a temporary shortage. Of course, new investments to better detect risks or to become more flexible and redundant have their own costs. We have to weigh the costs and benefits of "pay now or pay later."

Given the probability and severity of a risk, what can we do about it? As a first step, we need to know in general how we can manage risks. We turn to this next.



### 10.3.3 A Typology of Risk Management Methods

By definition, the way to manage a systematic risk is to diversify it away, as a financial portfolio manager might do by buying stocks from a variety of companies.

For total risk, there are a number of ways to manage risks:

- **Eliminate.** It is rare that one can entirely remove a risk, but one example would be to avoid doing something completely. Consider a manufacturer dealing with the risk of shipping delays arising from having a foreign supplier. For this example, we are not referring to the risk that the supplier does not produce on time, but to the risk that on-time productions are delayed by the long-distance shipper. We can eliminate the risk of such delays by relocating our manufacturing plant to the foreign supplier's neighborhood, thus eliminating the need for a long-distance shipper altogether.
- **Avoid.** More likely, we can avoid a risk by doing something else. For example, we can avoid the risk of delays from a foreign supplier by using a local supplier instead. The difference between "avoid" and "eliminate" is that in the former we change the source of the risk (the fact that the foreign supplier is located far away), while in the latter we do not.
- **Reduce.** To continue the example of risk of shipping delay from a foreign supplier, we could reduce the risk by using some backup, such as ensuring that the shipping company has spare capacity.
- **Transfer.** Purchasing insurance is a way to transfer a risk to another party. As another example, if a manufacturer's shipment from a foreign supplier is delayed, the manufacturer might pass that delay to its customer further down the supply chain, protected by a clause in its contract with that downstream customer. The difference between "reduce" and "transfer" in our shipping delay example is that in the former we deal with the risk at source (reducing risk at the foreign supplier), while in the latter we deal with it at destination (transferring the downside of a delay that has already happened to another party).
- **Transform.** One way to deal with a risk is to transform it into another risk that we can better accept or deal with. For example, one way to handle the risk of shipping delays from a foreign supplier is to have earlier forecasts, so that we can place our orders with the foreign supplier with a longer lead time. This transforms the risk of shipping delay to the risk of having the wrong forecasts.
- **Diversify.** We have considered diversification as the primary means of dealing with systematic risks. For example, we could diversify by having arrangements with alternate shippers. The distinction between "reduce" and "diversify" is that in the former we stick to the shipper in question, while in the latter we diversify away from that shipper.
- **Monitor.** Sometimes a risk is small enough that the best way to deal with it is to simply monitor it to ensure that it does not grow to become threatening. For example, suppose in

the worst case that a shipment does not even arrive; if we as a manufacturer suffer an amount that is acceptable in such an event, we would do best just to monitor the risk.

- **Accept.** It may be that the cost of managing a risk exceeds the benefit of doing so (e.g., the risk is unlikely to morph into a bigger risk); then we might as well do nothing and accept the risk as is.

With so many ways to deal with risks, what are we to do? This is where probability and severity become useful, as a risk map.

### 10.3.4 Risk Management with a Risk Map

Figure 10.3 shows a risk map, on which we can plot the probability and severity of each risk. The key point is that the risk map provides different action implications for different types of risks.

Those with high probability and high severity—at the top right of the risk map—require immediate diagnosis. Depending on the diagnosis, we should then neutralize the risk using almost any risk management method except to accept it.

Those with high probability but low severity—at the bottom right—may be monitored. We can also aim to reduce their high probabilities, and to avoid the risk altogether.

For risks with low probability but high severity—at the top left—it would make sense to explicitly prepare a response to their possible occurrence, by transforming the nature of the risks, transferring, or diversifying them away. It also makes sense to insure against such risks.

		Probability	
		Low	High
Severity	High	Transform, transfer, diversify, insure	Immediately diagnose and neutralize
	Low	Accept, insure	Monitor, reduce probability, avoid

Figure 10.3

Finally, risks that are of low probability and low severity—at the bottom left—should perhaps be accepted, as the cost of managing these risks might not be worth the benefit. At most, perhaps they should be transferred away, via insurance.

In plotting these risks on a risk map, it is useful to differentiate recurrent process risks—such as the risk of machine downtime in a production process—from one-off risks—such as the risk of a drug application being rejected by the FDA.

Importantly, these two types of risks have different value implications. As we show below, process risks usually affect value by reducing productivity. One-off risks usually affect value by directly reducing the value of a business. The two types also have different state-of-the-art tools to assess their riskiness. Therefore, we discuss them separately.

### 10.3.5 Advanced Topic: Process Risks

These risks are recurrent and routine. They involve processes in our business, such as purchasing, production, or forecasting.

Consider a prototypical production process with multiple stations. First, we analyze the probability and severity of risk—or downtime—at each station. These may be characterized by the mean time to failure (MTTF) and mean time to repair (MTTR). For example, if the MTTF of a station is 100 hours and the failure probability is constant over time, then the probability of failure is 1/100 over an hour. A long MTTF represents a low probability of failure, and a long MTTR represents a high severity of failure.

We now describe how MTTF and MTTR affect the value of the business in question, at the station and process level.

#### One Station

Consider first a station in the process. Suppose this station has a capacity of  $r$  units per hour if there were no risk of it being down. Now, with the MTTF and MTTR, its reduced capacity is:

$$r \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}} \quad (4)$$

That is,  $r$  is now scaled down by the station's availability, represented by the quotient. This, then, is the first performance effect of process risk: high probability (small MTTF) or high severity (big MTTR) leads to large reductions in capacity.

Process risk also has a second performance effect: it increases working capital in the form of inventory.

Consider figure 10.4. In panel (a), we have a station with no process risk—i.e., no breakdowns. Recall that it can produce units at a rate of  $r$  if there are no breakdowns. Suppose its inputs arrive at a rate of  $r_a$ . For simplicity, we consider the case when  $r_a$  is unvarying and is lower than  $r$ , so that the outflow rate—which is the minimum of  $r_a$  and  $r$ —is just  $r_a$ . In this case, the average inventory at the station is  $r_a/r$ .

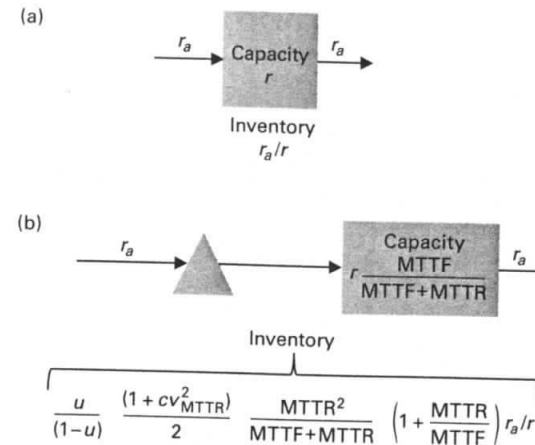


Figure 10.4

Process risk (variability) and working capital (inventory) at a station. No process risk (no downtime) (a); with process risk (MTTR, MTTF,  $cv^2_{\text{MTTR}}$ ) (b).

Now consider what happens when we have process risk, as in panel (b). As explained earlier, the first performance degradation is that the station has a reduced capacity, which is now  $r \times \text{MTTF}/(\text{MTTF} + \text{MTTR})$ .

Inventory at the station is now higher, at

$$r_a / \left( r \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}} \right),$$

or

$$\left( 1 + \frac{\text{MTTR}}{\text{MTTF}} \right) r_a / r.$$

Intuitively, a station with reduced capacity needs more time to process an input, so on the average the station now holds more inventory.

Worse, the MTTR itself is just a point estimate of process risk (it is the mean) from a probability distribution. Its variability may be measured by its *squared coefficient of variation* (denoted  $cv^2$ ), which is the variance divided by the square of the mean. Let us denote the  $cv^2$  of the MTTR by  $cv^2_{\text{MTTR}}$ . This variability  $cv^2_{\text{MTTR}}$  increases inventory even more.

To see this, note that a variable MTTR means that the processing time at the station is now variable, instead of being a constant. In queueing theory, the formula for the  $cv^2$  of a station's processing time is given by:

$$r(1 + cv_{MTTR}^2) \frac{MTTF}{\left(\frac{MTTF}{MTTR} + 1\right)^2}. \quad (5)$$

From queueing theory, we know that a variable processing time means that queues (i.e., inventory) will form when incoming units cannot be processed because they arrive when the station happens to be busy. Assuming that the incoming units themselves arrive with certainty, the average queue length is given by:

Station queue inventory =

$$\frac{u}{(1-u)} \times \text{station processing time} \times \frac{[cv^2 \text{ of station input arrivals}] + [cv^2 \text{ of station processing time}]}{2}, \quad (6)$$

where  $u$  is the utilization of the station. The utilization is also the average inventory in the station

$$\left(1 + \frac{MTTR}{MTTF}\right) r_a / r;$$

a high utilization is equivalent to high average inventory, since the unit stays in the station to be processed for much of the time. Returning to the queue inventory in equation (6), we can now substitute into it the station processing time, which is the reciprocal of equation (4). The  $cv^2$  of the input arrivals is 0, since by assumption, input arrivals do not vary. The  $cv^2$  of the station processing time is given in equation (5). Taken together, we get:

Station queue inventory =

$$\frac{u}{(1-u)} \frac{(1 + cv_{MTTR}^2)}{2} \frac{MTTR^2}{MTTF + MTTR}.$$

Therefore, as shown in figure 10.4, panel (b), not only do we have higher inventory at the station, we also have new inventory in a queue into the station. Both types of inventory—in the station as well as in the queue—increase with  $MTTR/MTTF$ . This makes intuitive sense, since a longer time for repairs compared to time to failure degrades performance. The inventory in the queue also increases with  $cv_{MTTR}^2$ , which again makes intuitive sense.

### Sequential Stations in a Process

Next we move from one station to a process with multiple stations. To keep the discussion manageable, suppose we have just two stations. We consider how performance is affected when the two stations are in sequence and when they are pooled in parallel.

Consider first a sequence of two stations, as in figure 10.5. As before, process risk has these effects:

- it reduces each station's capacity from  $r$  to

$$r \frac{MTTF}{MTTF + MTTR},$$

- it increases inventory at each station, from  $r_a/r$  to

$$\left(1 + \frac{MTTR}{MTTF}\right) r_a / r, \text{ and}$$

- it introduces inventory, in the form of queues, in front of each station.

What is different now is that Station 1's output is used as the input to Station 2, and because of that the input to Station 2 is varying—unlike that into Station 1, which is unvarying by assumption. Buzacott and Shanthikumar give a formula for the  $cv^2$  of the input into Station 2 as the product of  $u^2$  and the  $cv^2$  of Station 1's processing time from equation (5); that is:

$cv^2$  of Station 2 input arrivals =

$$u^2 r (1 + cv_{MTTR}^2) \frac{MTTF}{\left(\frac{MTTF}{MTTR} + 1\right)^2}. \quad (7)$$

This will in turn escalate the inventory in the queue in front of Station 2 further, to:

$$\frac{u}{(1-u)} \times \text{Station 2 processing time} \times \frac{[cv^2 \text{ of Station 2 input arrivals}] + [cv^2 \text{ of Station 2 processing time}]}{2}. \quad (8)$$

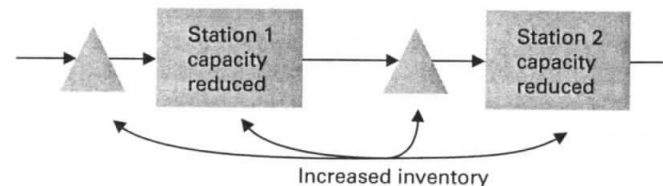


Figure 10.5  
Process risk for two stations in sequence.

The  $cv^2$  of Station 2's input arrivals is in equation (7), and the  $cv^2$  of Station 2's processing time is the same as that for Station 1, as given in equation (5).

Therefore, the Station 2 queue inventory is:

$$\frac{u}{(1-u)} \frac{MTTF + MTTR}{rMTTF} \left[ \frac{u^2 r (1 + cv_{MTTR}^2)}{\left(\frac{MTTF}{MTTR} + 1\right)^2} + \frac{r (1 + cv_{MTTR}^2)}{\left(\frac{MTTF}{MTTR} + 1\right)^2} \right] \\ = \frac{u(u^2 + 1)}{(1-u)} \frac{(1 + cv_{MTTR}^2)}{2} \frac{MTTR^2}{MTTF + MTTR}.$$

In other words, the queue inventory for Station 2 is amplified by  $(u^2 + 1)$  over that for Station 1. The higher Station 1's utilization  $u$ , the greater the amplification. At the extreme when Station 1's utilization is 1, Station 2's queue inventory is double that of Station 1.

#### Pooled Stations in a Process

Finally, we consider stations pooled together, as in figure 10.6.

How does process risk affect performance in this case? In some ways, the impact is similar to the previous case of sequential stations: (1) station capacities are reduced, (2) station inventories are increased, and (3) inventory appears in a queue before the stations. The difference here is the last. The formula for inventory in the queue when there are  $m$  pooled stations is:

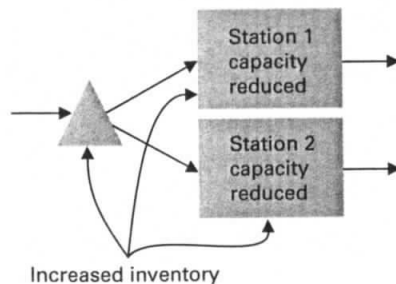


Figure 10.6  
Process risk for two stations pooled together.

$$\frac{u^{\sqrt{2(m+1)}-1}}{(1-u)} \times \frac{MTTF + MTTR}{mrMTTF} \times \frac{[cv^2 \text{ of station processing time}]}{2} = \\ \frac{u^{\sqrt{6}-1}}{(1-u)} \frac{(1 + cv_{MTTR}^2)}{4} \frac{MTTR^2}{MTTF + MTTR}. \quad (9)$$

While this is larger than the queue inventory for a single station as in equation (6), it is smaller than just the queue inventory for Station 2 in equation (8).

Taken together, this shows that process risk affects capacities and inventories. Unlike in finance, where risks are often thought of as pooled in a portfolio,<sup>2</sup> operational processes may amplify risks as they cascade through many stations in the processes.

#### Risk of Inadequate Warranty Reserves

In December 2006, Clay Sumner, an FBR Equity Research analyst, claimed that Dell Computers reported artificially high earnings by not setting aside adequate reserves to cover the warranty of computers sold. For many products from electronics to cars, the inadequate provision of warranty reserves is a major source of risk. How do we know how big the reserve ought to be? Intuitively, that should depend on:

- $S$  = the sales rate, or how many units we sell per year,
- $MTTF$  = mean time between failures, as above,
- $w$  = length of the warranty period (we assume this is exogenously given here; in section 15.4.3, we show how it is determined),
- $p$  = the unit price to the customer,
- $c$  = the unit cost to us,
- $t$  = time to when the unit fails, following its purchase at  $t = 0$ ,
- $b$  = our rebate to the customer when the unit fails during warranty.

There are two major types of warranties: (1) free replacement, in which  $b = c$ ; and (2) pro rata rebate, in which we give the customer a rebate of the price  $p$  she paid proportionate to the remaining life of the warranty; see figure 10.7.

The warranty reserve is then the expected cost of failures, or the sum (or integral, since we are dealing with continuous variables) of the failure probability and failure cost. The failure probability—assuming that failures are random—is given by the probability density function:

$$P_{(\text{unit fails at time } t)} = e^{-t/MTTF}/MTTF,$$

and the failure cost is  $\$b$  per unit, which depends on the warranty type: free replacement or pro rata. So for free replacement, the warranty reserve ought to be:



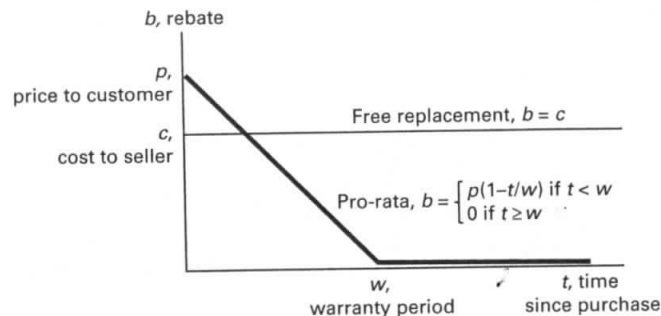


Figure 10.7  
Two types of warranties: free replacement and pro rata.

$$\int_0^w \left( \frac{e^{-\frac{t}{\text{MTTF}}}}{\text{MTTF}} \right) S c dt = S c \left( 1 - e^{-\frac{w}{\text{MTTF}}} \right)$$

and that for pro rata warranty ought to be:

$$\int_0^w \left( \frac{e^{-\frac{t}{\text{MTTF}}}}{\text{MTTF}} \right) \left[ S p \left( 1 - \frac{t}{w} \right) \right] dt = S p \left[ 1 - \frac{\text{MTTF}}{w} \left( 1 - e^{-\frac{w}{\text{MTTF}}} \right) \right].$$

In both cases, we see that as MTTF becomes larger—as failure becomes rarer—the appropriate warranty reserve goes down, as shown in figure 10.8 for an example in which  $S = 100$  units per year,  $c$  is \$1 per unit,  $p$  is \$1.2 per unit, and  $w$  is 5 years.

Finally, bear in mind that the reserves we just calculated are for one year. We should get the cumulative NPV of the reserves; for example, if the yearly reserves grow at  $g$  every year and should be discounted at  $r$ , the cumulative NPV is the yearly reserve divided by  $(r - g)$ .

### 10.3.6 One-off Risks and Value at Risk (VaR)

Like process risks, one-off risks can be evaluated on the basis of their probability and severity. But while the consequences of process risks tend to be gradual (e.g., lower capacity, higher inventory), the consequences of one-off risks tend to be categorical (e.g., patent application rejected, supply is tainted). In other words, one-off risks are less forgiving, and they require a different tool to evaluate their impact on valuation.

The most common tool comes under the value-at-risk (VaR) framework, in which we assess how much we would lose, given a probability and a time horizon. VaR is defined

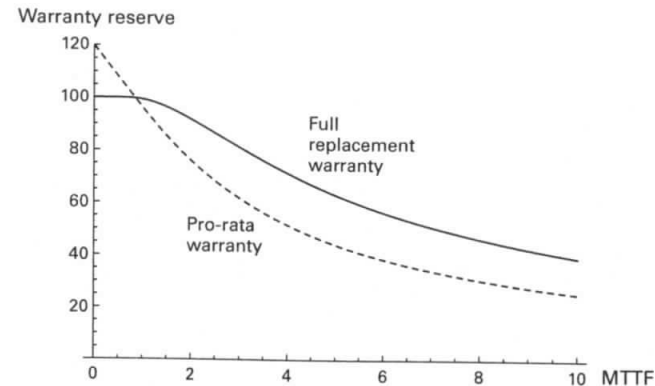


Figure 10.8  
Warranty reserves decrease with MTTF.

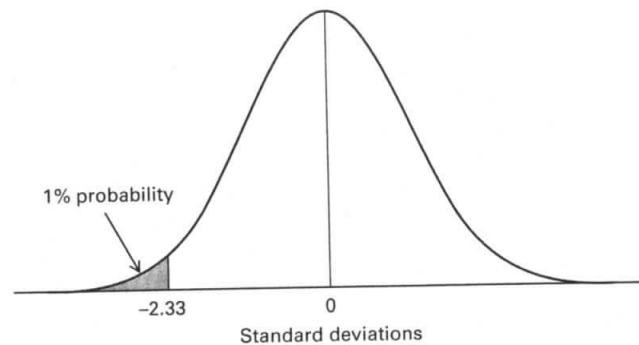
as the loss that would be incurred with some prespecified probability (say 1%) for a pre-specified portfolio (say \$100 million of real estate holdings). VaR was popularized by Banker's Trust and J. P. Morgan, the latter of which spun off a group called RiskMetrics to further develop the VaR tool.

Even though VaR was originally developed for assessing the risks of securities and credit portfolios, with some tweaking it offers a useful and simple method for assessing real assets. For example, in March 2000 thunderstorms hit Philips's semiconductor plant in Albuquerque, New Mexico. The plant furnace was hit by lightning, and the resulting fire destroyed the plant's clean room. The risk involved in this sort of one-off event could be quantifiable using the VaR methodology.

We describe two common methods for calculating VaR, then compare them.

### Parametric Method

The original RiskMetrics description of VaR uses historical data to estimate the parameters of an assumed normal probability distribution of financial assets. We can use it for real assets, although there are some necessary departures. First, the horizon for financial VaRs is usually in days, while that for real assets can reasonably be a lot longer. Suppose we want to find the 1%-probability-over-three-months VaR for a \$100 million portfolio of real risks: 40% of the value is in rentals for a clutch of retail stores and 60% in a three-month demand for our collection of skiwear products. We first assemble the historical three-month percentage changes for rentals and demand. The further back we go into the historical returns, the more confidence we have with our estimation of the probability distribution, *provided* the underlying distribution has not changed much. Using these



**Figure 10.9**  
Standard normal distribution.

historical percentage changes, we estimate the mean percentage change of the portion as  $40\% \times E(r_g) + 60\% \times E(r_m)$ , where  $E(\cdot)$  is the expectation operator to represent the mean, and  $r_g$  and  $r_m$  are the variables that represent the three-month percentage change of rental and demand. In some variants of this method, the percentage changes are the lowest within the three months. We calculate the variance of the portfolio acknowledging the covariance between the two stocks. The portfolio variance is  $40\%^2 \times \text{Var}(r_g) + 60\%^2 \times \text{Var}(r_m) + 2 \times 40\% \times 60\% \times \text{Cov}(r_g, r_m)$ , where  $\text{Var}(\cdot)$  is the variance operator and  $\text{Cov}(\cdot)$  the covariance operator.

Importantly, the parametric method assumes that the portfolio return is normally distributed with the estimated mean and variance. The 1% probability mark is 2.33 standard deviations lower than the zero mean in the standard normal distribution (see figure 10.9).

Let us suppose that the mean portfolio percentage change is 10% and the standard deviation (the square root of the variance) is 20%. Therefore, there is a 1% probability that the three-month percentage change is as much as  $10\% - 2.33 \times 20\% = -36.6\%$ . The \$100 million portfolio of risks will therefore have a VaR of  $\$100 \text{ million} \times 36.6\% = \$36.6 \text{ million}$ .

How realistic is this VaR? As we suggested, this depends on how representative the past is of the future. The more comprehensively we can model the risks, the more accurate can be our VaR. For example, we could account for changes in the mix of store locations, or changes in their retail formats. (As the clutch of rental locations expand from urban to rural areas, we should account for this change in mix.)

What about categorical risks like whether a patent application is approved? These VaRs tend to be volatile. For example, a patent valued at \$10 million may have a VaR of \$0 if the probability is 1%, but all \$10 million if the probability is (as an example) 0.1%. The

VaR amount is either \$0 or \$10 million, depending on what probability threshold we pick. In these cases, it makes sense to think through these risks differently. For example, we might focus on correlations—that is, if this patent application is rejected, will almost all related patent applications likewise be rejected? We might also think about divesting or hedging this sort of risk, perhaps through real options (see chapter 11).

### Historical Method

A major criticism of the parametric method is that it assumes a normal distribution. Simulation methods construct empirically based distributions instead. In this method, we use three-month changes in rentals and demands—going as far back as we might in the parametric method—to calculate the change in our portfolio. For example, for the three-month period of January-March 1983, rentals might have gone down 2% and demand 3%, and if this had been the case today, our portfolio would have gone down  $40\% \times 2\% + 60\% \times 3\% = 2.6\%$ .

This 2.6% downturn would constitute one of very many historical three-month changes, and we next put these changes into a histogram—rather than using them to estimate a normal distribution. Then the rest of the steps are as before: find the percent change corresponding to the 1% percentile in the probability distribution, and translate that percent change to a dollar change to get the VaR.

This simple historical method assumes that all historical changes are of equal weight. One modification is to place greater weight on more recent changes.

### Monte Carlo Method

A more sophisticated way is to explicitly model how values change over time, rather than using historical percentage changes.

Let us stay with our example of determining the 1%-probability-over-three-months VaR for a \$100 million portfolio of real risks: 40% of the value is in rentals for a clutch of retail stores and 60% in three months' demand for our collection of skiwear products.

We first build models of how rentals and demand might change. To do this, we postulate factors that drive these two risks. For example, we might have:

$$\text{rental}_{t+1} = \beta_1 + \beta_2 \times \text{rental}_t + \beta_3 \times t + \beta_4 \times \text{demand}_t + \varepsilon_t \quad (10)$$

$$\text{demand}_{t+1} = \alpha_1 + \alpha_2 \times \text{demand}_t + \alpha_3 \times \text{SKU}_t + \alpha_4 \times \text{rental}_t + \eta_t \quad (11)$$

The  $t$  subscripts indicate three-month periods. In other words, rental in the next three months depends on rental in this three-month period, a trend term (the  $t$  term), demand in this period, and a random noise term with zero mean and a normal distribution. Demand in the next three-month period depends on demand in the current period, the number of stock-keeping units (SKUs) in this period, the rental in this period (which also provides a

proxy for the total retail square footage), and another random noise term, also with zero mean and a normal distribution. Notice how we connect rental and demand in the two equations, so as capture the correlation between the two.

Using historical data containing observations of rental, inflation, demand, stores, and number of SKUs for three-month periods, we can run regressions to get the  $\beta$  and  $\alpha$  terms. The regression results also produce the variance for  $\varepsilon$  and  $\eta$ .

Now we create a distribution of the value of the portfolio consisting of 40% of its value in rentals for a clutch of retail stores and 60% in skiwear demand. To do this, we plug in what we now know for all the values on the right-hand side of equations (10) and (11), as well as randomly generate values of  $\varepsilon$  and  $\eta$  based on normal distributions with their variances estimated from the regressions. Then we put these generated portfolio values in a histogram to find the value corresponding to the 1% percentile in the probability distribution. The difference between this value and the current value is the VaR.

Of course, we can add more realism—and complexity—to the procedure just described. Importantly, we have assumed that the connection between rental and demand is static, and that the stochastic components  $\varepsilon$  and  $\eta$  are independent. We could add more realism by explicitly correlating  $\varepsilon$  and  $\eta$  with, say, a bivariate normal distribution. Also, we actually make only one forecast, so all the histogram is showing is the randomness in  $\varepsilon$  and  $\eta$ . We could use short historical periods—say one month rather than three months—and use a sequence of one-month predictions to get our three-month VaR. We could also add more explanatory factors other than just a trend term or the number of SKUs. And we could assume that  $\varepsilon$  and  $\eta$  do not follow the normal distributions.

### Comparison and Critique

We can compare the methods along several dimensions, and suggest situations in which one method might be superior to another:

- **Distributional assumptions.** The parametric method assumes that the probability distribution is normal, and can be completely parameterized by the mean and variance. The historical and Monte Carlo methods escape this criticism. But like the parametric method, the historical method still assumes that the past distribution is representative of the future. If we are dealing with very liquid risks—such as those in the pricing of financial assets—the normality assumption under the parametric method seems reasonable. At the opposite extreme, if we are dealing with rare risks, then the Monte Carlo method seems to be appropriate, since in principle it simulates as many different scenarios as are needed. Nevertheless, the Monte Carlo method is still susceptible to “model risk”—that is, the model used deviates too much from the actual distribution. It also assumes that we know the probability distributions of the risk factors.
- **Computational requirements.** Not surprisingly, the parametric method—which requires only two parameters—is computationally the easiest. At the opposite extreme is

the Monte Carlo method, which is known to take several orders of magnitude more of computational power and time.

Apart from the problems just indicated, the VaR methodology in general is subject to many other criticisms. For one, it provides only a threshold loss amount. A VaR amount of \$1 million with 1% probability could also represent a risk of losing \$100 million with 0.9% probability. For another, by providing just a single statistic, it could offer a false sense of security in a black box. The Monte Carlo regressions in which we can explicitly model what is driving changes in value are very helpful, but they are only as helpful as the models themselves. Using VaR does not make us less susceptible to risks we do not know about. And finally, the provision of a single statistic also subjects it to manipulation, as when managers decide to provide “the number that the CEO likes to see.” But many of these criticisms can also be leveled against most other measures of risk, and the VaR has withstood several decades of practical use in industry.

### 10.4 Takeaways and Toolkit

In this chapter, we learn why we have to dig below financial statements to better value a company's risks. To do that, we need to:

- Differentiate between systematic and total risk. If our risks can be diversified away, then it is useful to think only about systematic risk, which is the undiversifiable part of total risk.
- To find systematic risk, we calculate beta. Operations affect beta through operational leverage: the higher the proportion of costs that are fixed, the higher the beta.
- To find total risk, we first have to specify the myriad of risks that might occur. A risk map of probabilities and severity can help. In finance, probabilities (or variances) may be reduced through diversification. That is true in operations; but in operations, probabilities could also be amplified.
- It is also helpful to analyze two types of total risk: process and one-off risks. They affect firm value differently, and are amenable to different tools for assessment.
- Process risk reduces capacity and increases working capital (inventory). One-off risks can be assessed using a value-at-risk framework.

In table 10.1, we summarize the formulae for systematic and total risk.

### 10.5 Survey of Prior Research in One Paragraph

The leading researchers who have developed the idea of beta include Harry Markowitz, Bill Sharpe, John Lintner, and Jack Treynor. Recent research argues that specific risks are

**Table 10.1**  
Toolkit: Formulae for Systematic and Total Risk

Type of Risk	Formula
Systematic risk	$\text{beta} = \text{Cov}(r_{\text{business}}, r_{\text{market}}) / \text{Var}(r_{\text{market}})$ $= \left[ 1 + \frac{\text{PV}(\text{fixed cost})}{\text{PV}(\text{asset})} \right] \text{beta}_{\text{revenue}}$
Total process risk	<p>Inventory needed to buffer process risks (captured by MTTF and MTTR):</p> <ul style="list-style-type: none"> <li>One station: <math display="block">\frac{u}{(1-u)} \frac{(1 + cv_{\text{MTTR}}^2)}{2} \frac{\text{MTTR}^2}{\text{MTTF} + \text{MTTR}} + \left( 1 + \frac{\text{MTTR}}{\text{MTTF}} \right) r_a / r</math> </li> <li>Two stations in a sequence: as above for first station plus below for second station: <math display="block">\frac{u(u^2 + 1)}{(1-u)} \frac{(1 + cv_{\text{MTTR}}^2)}{2} \frac{\text{MTTR}^2}{\text{MTTF} + \text{MTTR}} + \left( 1 + \frac{\text{MTTR}}{\text{MTTF}} \right) r_a / r</math> </li> <li>Two stations in a pool: <math display="block">\frac{u^{\sqrt{6}-1}}{(1-u)} \frac{(1 + cv_{\text{MTTR}}^2)}{4} \frac{\text{MTTR}^2}{\text{MTTF} + \text{MTTR}} + 2 \left( 1 + \frac{\text{MTTR}}{\text{MTTF}} \right) r_a / r</math> </li> </ul>
Total one-off risk	<p>Calculate the value at risk (VaR) using one of these methods:</p> <ol style="list-style-type: none"> <li>1. Parametric</li> <li>2. Historical</li> <li>3. Monte Carlo</li> </ol>

not completely diversifiable. Two leading finance scholars are Andrei Shleifer and Robert Vishny. Richard Bettis makes a similar point, from the strategy angle, and John Birge and Paul Glasserman are leading scholars in applying operations research techniques to risk assessment. The treatment of warranty reserves here is by Warren Menke. Kevin Hendricks and Vinod Singhal pioneer the quantification of operational disruptions on firms' financial performance. Marshall Fisher, Jan Hammond, and Ananth Raman have done rigorous analyses on collective forecasting and the staging of production capacities to increase flexibility, so as to mitigate the risk of demand-supply mismatches. Yossi Sheffi has written a useful reference called *The Resilient Enterprise*. Value at risk has long been known in statistics, but its current form is popularized by J. P. Morgan under the name RiskMetrics™.

## 10.6 Further Reading

- Bettis, R. A. 1983. "Modern Financial Theory, Corporate Strategy and Public Policy: Three Conundrums." *Academy of Management Review* 8 (3): 406–415.
- Brealey, R. A., S. C. Myers, and F. Allen. 2006. *Principles of Corporate Finance*. Boston: McGraw-Hill/Irwin.
- Buzacott, J. A., and J. G. Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems*. Englewood Cliffs, NJ: Prentice-Hall.
- Foster, D. P., and R. A. Stine. 2006. "Being Warren Buffett: A Classroom Simulation of Risk and Wealth When Investing in the Stock Market." *American Statistician* 60: 53–60.

- Lintner, J. 1965. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics*: 13–37.
- Markowitz, H. 1952. "Portfolio Selection." *Journal of Finance*: 77–91.
- Menke, W. W. 1969. "Determination of Warranty Reserves." *Management Science* 15 (10): 542–549.
- Morgan, J. 1995. *Riskmetrics Technical Manual*. New York: J. P. Morgan.
- Sharpe, W. F. 1963. "A Simplified Model for Portfolio Analysis." *Management Science* 9 (2): 277–293.
- Sheffi, Yossi. 2007. *The Resilient Enterprise*. Cambridge, MA: MIT Press.
- Treynor, J. 1961. "Toward a Theory of Market Value of Risky Assets." Unpublished manuscript.