Financial Intermediation and the Price of Risk

Hyun Song Shin Princeton University

"Global Financial Crisis of 2007 – 2009:
Theoretical and Empirical Perspectives"

Summer Economics at SNU and Korea Economic Association
Seoul National University, 2009

Issues to be Addressed

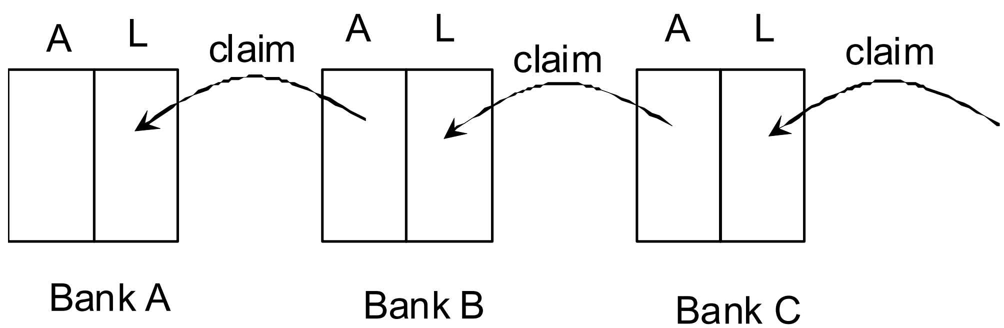
- What are the fundamental causes of boom-bust cycles?

- How can we make sense of the "underpricing of risk"?

- What role does *actual defaults* play, as compared to *potential defaults*?

- What is the role of the structure of the financial system in exacerbating boom-bust cycles?

Domino Hypothesis



- Claim is that the channel of financial contagion is through **chain of defaults.**
  - Passive players, who stand by while others fail
  - No role for prices
  - Only implausibly large shocks generate any contagion in simulations

How plausible?

Two issues:

- Defaults
- Chain

Return later to address these questions.

In 2007/8 crisis, direction of contagion has been reversed.

Bear Stearns, Lehman Brothers and Northern Rock crises were **runs** on the liabilities side.

Value-at-Risk

Value at Risk

Defining and managing risk is one of the most important issues facing firms in their daily operations, and especially important for financial institutions that rely on leverage.

Informally speaking, the concept of *value at risk* can be motivated by the need to find an answer to the following question:

> "*What (realistically) is the worst that could happen over one day, one week, or one year?*"

We can imagine all kinds of bad outcomes that may conceivably happen, but running a firm or a bank with such worst case scenarios in mind would be pretty debilitating.

Instead, the important qualifier in the question above is that the worst case outcome is something that we could *realistically* expect to happen. The question is how we should define "realistically".

If we had some idea of the probabilities with which the possible outcomes transpire, we could try to put some numerical magnitudes on what we mean by "realistically".

Value at risk (VaR) is an answer to the question above where "realistically" is defined by finding an outcome that is so bad that anything worse is highly unlikely. More precisely, value at risk is the realistically worst case outcome in the sense that anything worse only happens with probability less than some fixed level (such as 1%).

Definition. Let W be a random variable. The **value at risk** at confidence level c relative to base level W_0 is the smallest non-negative number denoted by VaR such that

$$\text{Prob}\left(W < W_0 - \text{VaR}\right) \leq 1 - c$$

The random variable W could be, for instance, the market value of a portfolio at the end of the year. In other words, the Value-at-Risk is a quantile of the loss distribution, where the loss is measured from the base level W_0. Denoting by F the cumulative distribution function of W, the Value-at-Risk is given by

$$\text{VaR } = \inf\left\{V | F\left(W_0 - V\right) \leq 1 - \alpha\right\} \tag{1}$$

Notice that the definition of value at risk specifies three things.

- Some random variable W. By specifying a random variable, we are already implicitly defining

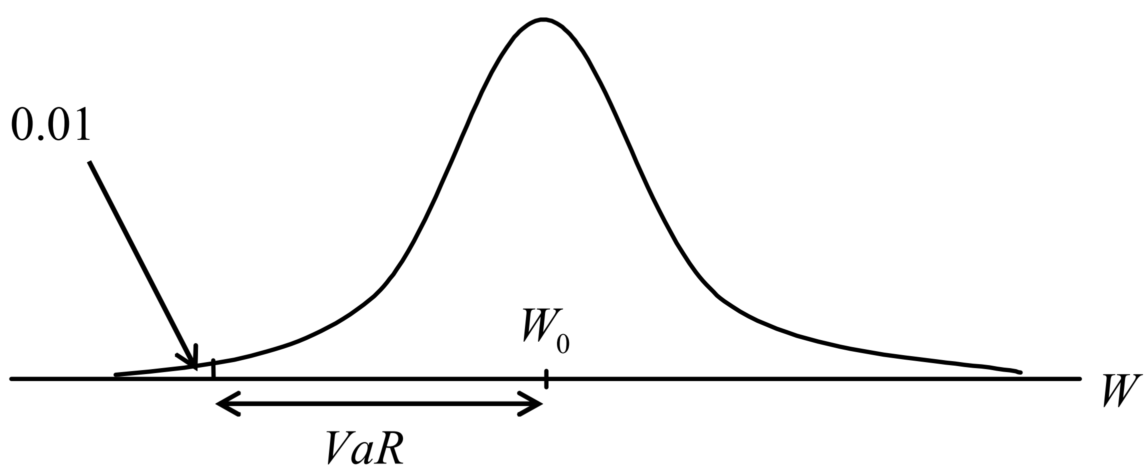
- the *time horizon* and
- the process governing the evolution of W over time.

- Some base level W_0 from which final outcomes can be measured

- Confidence level c which gives concrete meaning to "realistically worst case" outcome.

Example. The random variable W is the value of a portfolio in three months' time. W has a known probability density, as depicted in the figure below. The base level W_0 is today's value of the portfolio, and so outcomes are measured relative to today's value. The confidence level c is set at 99%.

The value at risk over the next three month interval relative to today's value at confidence level 99% is the distance indicated by the double arrow.



Rationale for Value at Risk

In what sense is value at risk the right measure of risk? There are two broad rationalizations of VaR that are typically given.

The first is as a **measure of potential extreme loss**. But in this case, VaR is not ideal. VaR is a worst case scenario, where anything worse happens with probability less than $1 - c$. However, conditional on something really bad happening, we would like to know how bad things can get. Value at risk is not good at giving that kind of measure.

Example. A bank is owed \$1 million by firm A, and is owed \$1 million by firm B. Both firms are creditworthy and will pay the \$1 million with probability 0.995. However, when firm A defaults, the bank still recovers \$0.5 million by selling the collateral put up by firm A. But when firm B defaults, the bank recovers nothing. The VaR relative to \$1 million at 99% confidence is zero in both case.

Why? Let's go to the definition of VaR. We want the smallest non-negative number x for which

$$\text{Prob}\left(\text{repayment} < \$1m - x\right) \leq 0.01 \tag{*}$$

But since both firms will repay with probability 0.995, we can set $x = 0$, and still have

$$\text{Prob}\left(\text{repayment} < \$1m\right) = 0.005 \leq 0.01$$

Thus, (*) holds with $x = 0$, and so value at risk is zero. This is true for both firm A and firm B. However, the bank has more to lose when firm B defaults than when firm A defaults.

If what you are interested in is the size of the potential loss conditional on something really bad happening, then value at risk is not a good measure.

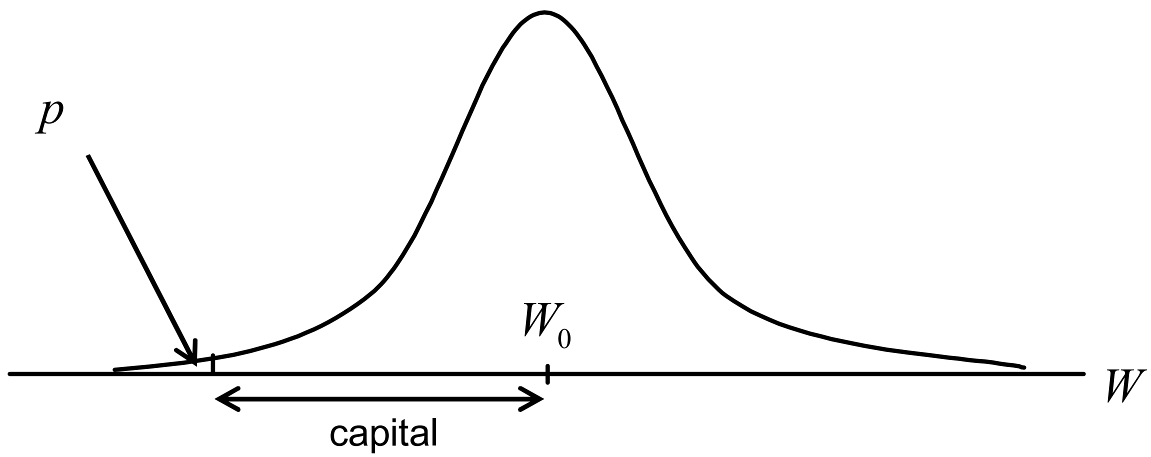
A better measure for the conditional loss would be the *conditional expected loss*, sometimes known as **tail loss**. The tail loss is defined as the expected loss conditional on the random variable W falling below some threshold point q.

$$\text{Tail loss } = \mathsf{E}\left(W|W<q\right) = \frac{\int_{-\infty}^{q} W f\left(W\right) dW}{\int_{-\infty}^{q} f\left(W\right) dW}$$

The second rationale for the concept of value at risk is in terms of the **capital** that a firm or bank must hold against possible failure, and for internal control purposes.

See the figure below.

Suppose W measures the value of total assets of the firm at some fixed date in the future, and W_0 is today's value of the firm's assets. If the firm has capital (equity) indicated by the size of the double arrow, then the firm will remain solvent as long as W does not fall below W_0 − capital. So, the probability that the firm will go bankrupt is p. By holding larger capital, the probability of failure can be decreased further.



If the firm or bank has limited liability, then it does not matter whether the firm goes bust just marginally, or whether it goes bust spectacularly, leaving a big shortfall. In this sense, the tail loss is not a concern for a firm with limited liability.

For such a firm, value at risk gives the correct measure of the amount of capital that it must hold so as to stave off default.

This is related to our earlier discussion on the Modigliani and Miller theorem, where we noted the divergent interests of debt holders and equity holders with regard to the riskiness of the firm's assets.

In addition, if the firm gets into difficulties, there may still be time to recover the situation by taking corrective action. The fact that value at risk is measured with a fixed horizon means that the time dimension can be set in a reasonable way for the particular case at hand.

The idea is that the time horizon in the definition of VaR is long enough so that

- corrective measures can be taken to rectify the problem (e.g. through recapitalization, rescue, reorganization)

- the horizon is also appropriate for the degree of illiquidity of the assets held by the firm

- confidence level chosen based on ease with which potential investors can be recruited in the recapitalization

To summarize, value at risk is a better measure of risk when it is interpreted as a buffer against possible failure. For this reason, value at risk plays an important role in *financial regulation.*

1988 Basel Capital Accord, and 1996 Amendment for Market Risk

The Basel Capital Accord is an agreement reached in 1998 among the member countries of the Basel Committee for Banking Supervision to govern the regulation of so-called “internationally active” banks. The main part of the 1988 accord concerns capital that banks are required to hold against the possible losses from defaulting borrowers.

The 1996 amendment to the Basel Accord required banks to hold capital against **market risk** - that is, the loss resulting from price changes of securities and other traded assets. The required capital is based on a value at risk concept with

- horizon of 10 trading days or two calendar weeks

- 99% confidence level

- observation period based on at least a year of data, updated at least once a quarter

- Capital is higher of:
  - previous day's VaR
  - average VaR over last 60 business days times a "multiplier" k. Multiplier $k \geq 3$. Higher k imposed by local regulators if back-testing reveals large number of exceptions.

We will return to review a possible reason for having a multiplier of 3.The concept of *Value-at-Risk* is motivated by the need to find an answer to the following question:

> "*What (realistically) is the worst that could happen over one day, one week, or one year?*"

Portfolio Choice

Portfolio decisions when an investor uses Value-at-Risk to manage risk.

Two assets - a risky security and cash.

The price of the risky security at date t is p_t, and number of units of the risky security is y_t.

Cash at date t is c_t.

Price of the risky security next period (at date $t+1$) is p_{t+1}, and is uncertain when viewed from date t.

$\tilde{r}_{t+1}$ the return from date t to date $t+1$ on the risky asset. Then

$$p_{t+1} = (1 + \tilde{r}_{t+1})\, p_t \tag{2}$$

$\tilde{r}_{t+1}$ is independent across dates, and is identically distributed at all dates with mean $\mu > 0$ and variance σ^2.

The capital of the investor at date t is denoted by e_t.

The investor who borrows in order to buy more of the risky asset has a balance sheet at date t which could be depicted as follows. The investor incurs debt of $-c_t > 0$ and buys risky securities worth the sum of his own capital e_t and the borrowed money $-c_t$.

Assets	Liabilities
Securities $p_t y_t$	Equity e_t Debt $-c_t$

(3)

The leverage of this investor is given by

$$\frac{p_t y_t}{e_t} \tag{4}$$

The balance sheet of this pessimistic investor can be depicted as follows.

Assets	Liabilities
Cash c_t	Equity e_t Securities $-p_t y_t$

(5)

The leverage of the pessimistic investor with balance sheet (5) is

$$\frac{e_t - p_t y_t}{e_t} \tag{6}$$

Long-only investor:

Assets	Liabilities
Cash c_t Securities $p_t y_t$	Equity e_t

(7)

Whether (3), (5) and (7) balance sheet identity holds at every date t.

$$p_t y_t + c_t = e_t \tag{8}$$

The new value of capital e_{t+1} is a function of the new realized price p_{t+1}, and satisfies:

$$\begin{aligned}
e_{t+1} &= p_{t+1}y_t + c_t \\
&= p_{t+1}y_t + e_t - p_t y_t \\
&= (p_{t+1} - p_t)\, y_t + e_t \\
&= \left[(1 + \tilde{r}_{t+1})\, p_t - p_t\right] y_t + e_t \\
&= \tilde{r}_{t+1} p_t y_t + e_t
\end{aligned} \tag{9}$$

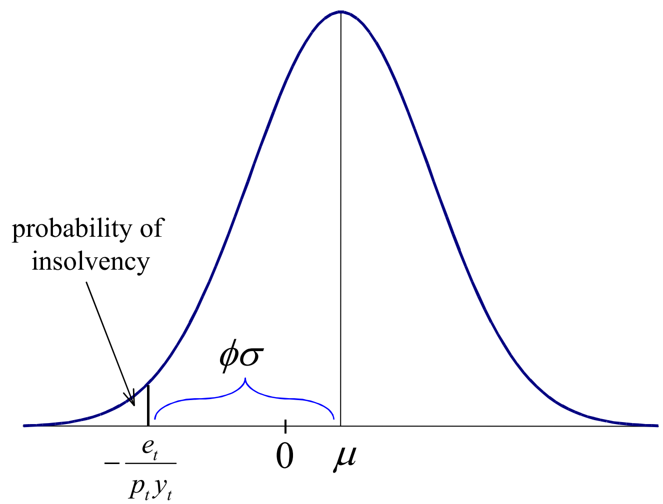


Figure 1: Probability density of $\tilde{r}_{t+1}$

The investor becomes insolvent if $e_{t+1} \leq 0$. From equation (9), this happens when the return on the risky security is sufficiently bad so that $\tilde{r}_{t+1} p_t y_t + e_t \leq 0$, or

$$\tilde{r}_{t+1} \leq -\frac{e_t}{p_t y_t} \tag{10}$$

The smaller is the initial equity level e_t or the larger is the initial holding y_t, the greater is the chance of going bust. Let ϕ be defined as the constant for which we have

$$\text{Prob } (\tilde{r}_{t+1} \leq \mu - \phi\sigma) = 1 - \alpha \tag{11}$$

In other words, $\phi\sigma$ is the Value-at-Risk for the risky return $\tilde{r}_{t+1}$ at the confidence level α relative to the mean return μ. Then, by choosing the size of the holding of the risky asset y_t, the investor can ensure that the probability of his becoming insolvent next period is kept at most $1 - \alpha$. From Figure 1 we see that the probability of insolvency is exactly $1 - \alpha$ when

$$\mu + \frac{e_t}{p_t y_t} = \phi\sigma \tag{12}$$

Solving for the dollar value of the risky security position, we have

$$p_t y_t = \frac{e_t}{\phi\sigma - \mu} \tag{13}$$

The investor cannot hold any more than this amount of the risky security, since then the probability of insolvency rises above the threshold value $1-\alpha$, thereby violating his Value-at-Risk Constraint.

Will the investor hold any less than the amount in (13)? No, since we are considering the case where $\mu > 0$ so that from (9), we have

$$E\left(e_{t+1}\right) = \mu p_t y_t + e_t \tag{14}$$

The expected equity value next period is strictly increasing in y_t, so that the investor wishes to hold as much of the risky security as is permitted by his Value-at-Risk constraint. This is a consequence of the fact that the return on the risky security is strictly higher than that on cash. The upshot is that the investor's holding of the risky security is given exactly by (13).

Upward-Sloping Demand Reactions

Another perspective on the investor's decision is to consider the leverage maintained on the balance sheet. From equation (13) we see that the investor's leverage is given by

$$L = \frac{p_t y_t}{e_t} = \frac{1}{\phi\sigma - \mu} \tag{15}$$

Given our assumption of constant μ and σ, leverage is also constant. Therefore, one way we can characterise the investor's portfolio decision is one of maintaining constant leverage in the face of price changes. However, leveraging targeting entails upward-sloping demand responses and downward-sloping supply responses - that is, the investor will buy more of the risky security if its price rises, and sells some of the risky security if the price falls. Such price responses provide the pre-conditions for amplifications of shocks.

In order to appreciate the consequences of a leverage target on the demand and supply responses to price changes, let us first consider a simple numerical example of an investor who aims to maintain a *constant* leverage of 10. The initial balance sheet is as follows. The investor holds 100 dollars worth of securities, and has funded this holding with debt worth 90.

Assets	Liabilities
Securities, 100	Equity, 10
	Debt, 90

Assume that the price of debt is approxminately constant for small changes in the price of the securities, so that the burden of adjustment falls on the equity. Suppose the price of securities increases by 1% so that the dollar holding of the securities rises to 101.

Assets	Liabilities
Securities, 101	Equity, 11
	Debt, 90

At this new higher price, the equity rises to 11, so that leverage then falls to $101/11 = 9.18$. This is because the equity rises by a much larger percentage rate (10%) due to the leverage. At the higher level of equity, the investor can restore leverage by taking on additional debt of D to purchase D worth of securities on the asset side so that

$$\frac{\text{assets}}{\text{equity}} = \frac{101 + D}{11} = 10$$

The solution is $D = 9$. The investor takes on additional debt worth 9, and with this money purchases securities worth 9. Thus, an increase in the price of the security of 1 leads to an increased holding worth 9. The demand response is upward-sloping. After the purchase, leverage is now back up to 10.

Assets	Liabilities
Securities, 110	Equity, 11
	Debt, 99

The mechanism works in reverse, too. Suppose there is shock to the securities price so that the value of security holdings falls to 109. On the liabilities side, it is equity that bears the burden of adjustment, since the value of debt stays approximately constant.

Assets	Liabilities
Securities, 109	Equity, 10
	Debt, 99

Leverage is now too high $(109/10 = 10.9)$. The investor can adjust down his leverage by selling securities worth 9, and paying down 9 worth of debt. Thus, a *fall* in the price of securities of leads to *sales* of securities. The supply response is downward-sloping. The new balance sheet then looks as follows. The balance sheet is now back to where it started before the price changes. Leverage is back down to the target level of 10.

Assets	Liabilities
Securities, 100	Equity, 10
	Debt, 90

In constrast to the textbook demand and supply responses to price changes, we see that investors with Value-at-Risk constraints exhibit perverse demand and supply responses where higher prices lead to purchases and lower prices lead to sales. To see how the magnitudes relate to the leverage targeted by the investor, note from (9) and (13) that the proportion change in equity can be written as

$$\frac{e_{t+1} - e_t}{e_t} = \tilde{r}_{t+1} \frac{p_t y_t}{e_t} = \tilde{r}_{t+1} \cdot L \tag{16}$$

while the proportional change in total assets as a consequence of the price change (but before the portfolio adjustment) is

$$\frac{p_{t+1} y_t - p_t y_t}{p_t y_t} = \tilde{r}_{t+1} \tag{17}$$

Comparing (16) and (17), we see that for a leveraged investor, equity rises

L-times faster than total assets. The price response of the investor can be obtained by tracking the new quantity y_{t+1}. Since the investor maintains constant leverage L, we have

$$\frac{p_t y_t}{e_t} = \frac{p_{t+1} y_{t+1}}{e_{t+1}} = L \tag{18}$$

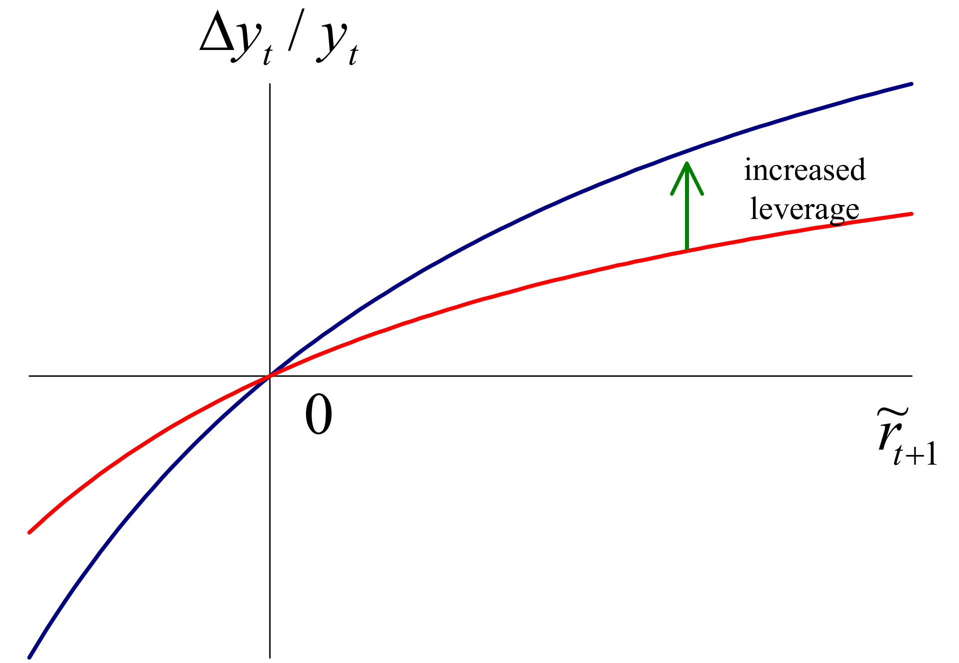
Hence

$$\frac{y_{t+1}}{y_t} = \frac{e_{t+1}/e_t}{p_{t+1}/p_t} = \frac{1 + \tilde{r}_{t+1} \cdot L}{1 + \tilde{r}_{t+1}} \tag{19}$$

The proportional increase in the holding of the risky security can be expressed as a function of the return on the risky asset $\tilde{r}_{t+1}$ and the degree of leverage L.

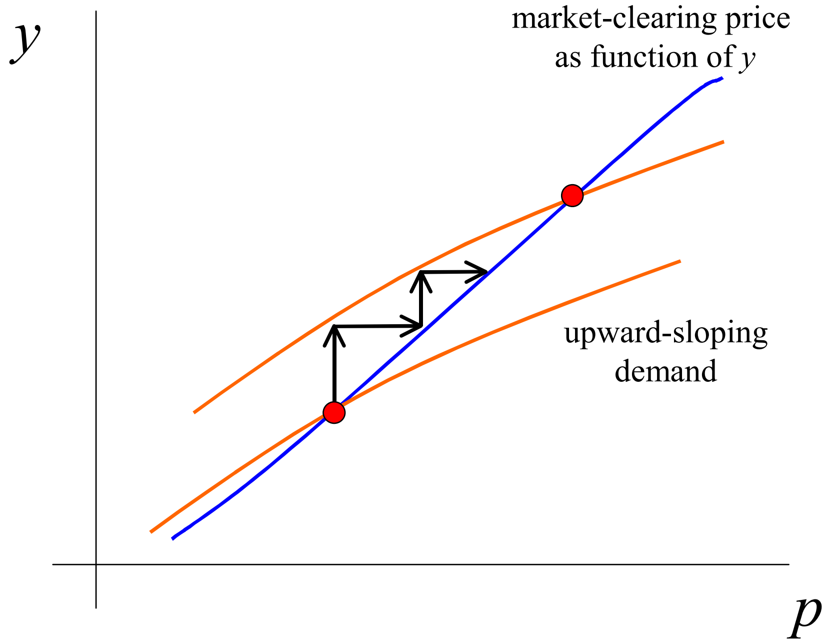
$$\frac{y_{t+1} - y_t}{y_t} = \frac{\tilde{r}_{t+1}}{1 + \tilde{r}_{t+1}} \cdot (L - 1) \tag{20}$$

The price response is upward-sloping in the return $\tilde{r}_{t+1}$, and is illustrated in Figure **??**. The higher is the target leverage L maintained by the investor, the steeper is the demand response to price changes.

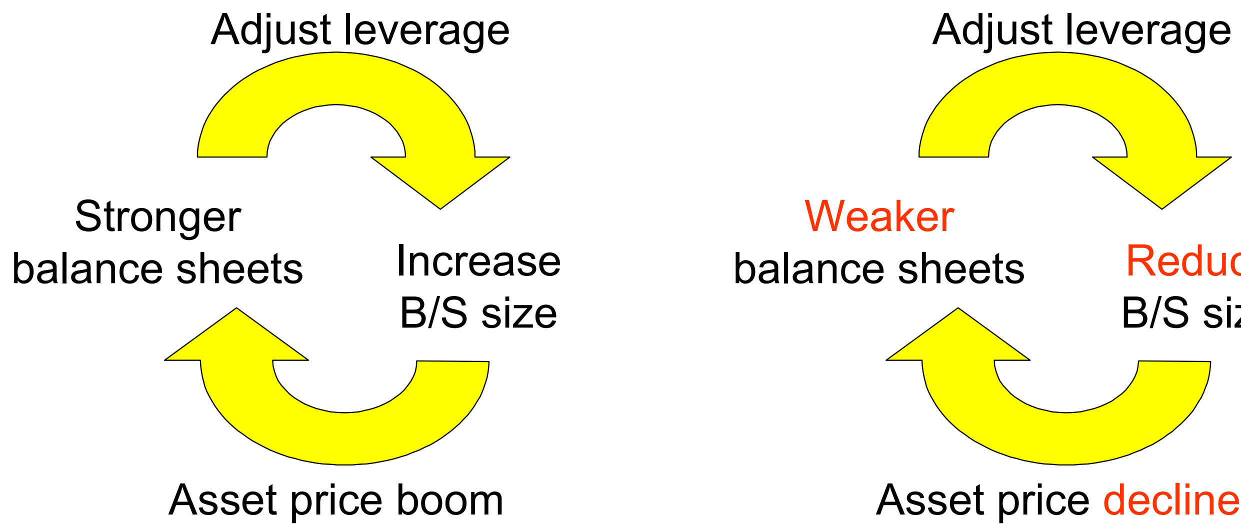


Upward-sloping demand response to $\tilde{r}_{t+1}$

Feedback



Amplification cycle of capital enhancement and increase in price



Feedback "on the way up" Feedback "on the way down"

General Equilibrium with Value-at-Risk

Two key ingredients:

- No defaults
- No chains among banks

In practice, both defaults and chains are important.

But purpose of these assumptions is to emphasize that the fundamental source of financial fluctuations is balance sheet dynamics.

Therefore we examine a simplified model with no default and no chains among banks.

General Equilibrium with Value-at-Risk

Two dates, 0 and 1.

Single risky security and cash

Risky security's payoff is random variable $\tilde{w}$, with expected value $q > 0$.

Random variable $\tilde{w}$ is uniformly distributed over the interval:

$$[q - z, q + z]$$

$z > 0$ is a known constant.

Mean and variance of $\tilde{w}$ are

$$\begin{aligned} E(\tilde{w}) &= q \\ \sigma^2 &= \frac{z^2}{3} \end{aligned}$$

Cash pays interest rate of zero.

p is price of the risky security.

For investor with equity e who holds y units of the risky security, payoff of the portfolio is the random variable:

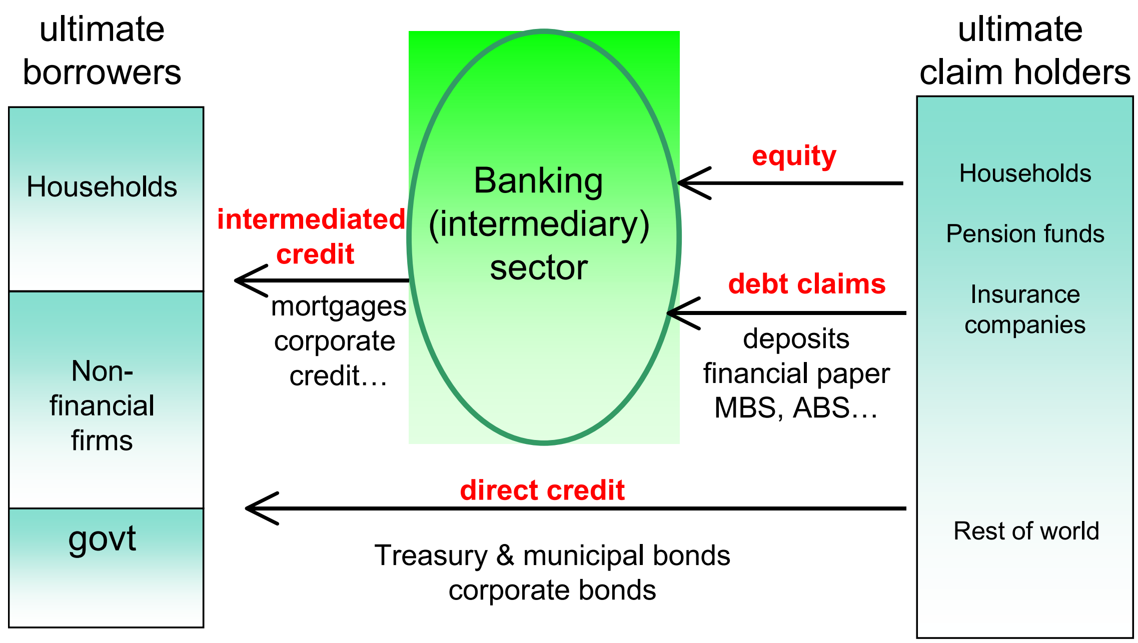
$$W \equiv \tilde{w}y + (e - py) \qquad (21)$$

Two groups of investors - *passive investors* and *active investors*.

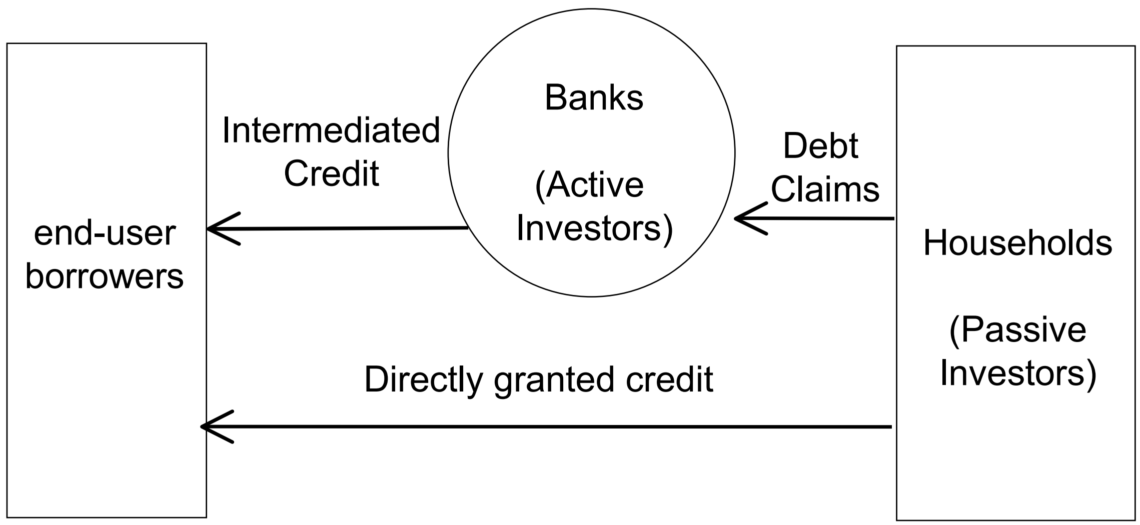
Interpetation

- Risky securities are *loans* granted to ultimate borrowers
- Market value risky security is market value of loans.
- Passive investors' holding of risky security is credit granted *directly* by the household sector (e.g. holding of corporate bonds)
- Active investors's holding of risky security is *intermediated finance*: active investors are banks who borrow from the households in order to lend to the ultimate borrowers.

Stylised Financial System



Simplified Financial System



Intermediated and Directly Granted Credit

Passive investors have mean-variance preferences over the payoff from the portfolio. They maximise

$$U = E(W) - \frac{1}{2\tau}\sigma_W^2 \tag{22}$$

$\tau > 0$ is constant called the investor's "risk tolerance" and σ_W^2 is the variance of W. In terms of the decision variable y, passive investor's objective function is

$$U(y) = qy + (e - py) - \frac{1}{6\tau}y^2z^2 \tag{23}$$

Optimal holding of risky security satisfies first order condtion:

$$q - p - \frac{1}{3\tau}z^2y = 0 \tag{24}$$

Optimal risky security holding of the passive investor (denoted by y_P) is

$$y_P = \begin{cases} \dfrac{3\tau}{z^2}\,(q-p) & \text{if } q > p \\ \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

These linear demands can be summed to give the **aggregate demand**.

If τ_i is the risk tolerance of the ith investor and $\tau = \sum_i \tau_i$, then (25) gives the aggregate demand of the passive investor sector as a whole.

Portfolio decision of the active (leveraged) investors.

Active investors are risk-neutral but face Value-at-Risk (VaR) constraint.

General VaR constraint is that the capital cushion be large enough that the default probability is kept below some benchmark level. Consider special case where that benchmark level is zero. Then, the VaR is the condition that leveraged investors issue only risk-free debt.

The constraint is that the investor's capital (equity) e be large enough to cover this Value-at-Risk. The optimization problem for an active investor is:

$$\max_{y} E(W) \quad \text{subject to VaR} \leq e \tag{26}$$

If the price is too high (i.e. when $p > q$) the investor holds no risky securities.

When $p < q$, then $E(W)$ is strictly increasing in y, and so the Value-at-Risk constraint binds.

Optimal holding of the risky security satisfies VaR $= e$.

Assets	Liabilities
securities, py	equity, e debt, $py - e$

Value-at-Risk constraint stipulates that the debt issued by the investor be risk-free.

For each unit of security, minimum payoff is $q - z$. In order for the investor's debt to be risk-free, y should satisfy $py - e \leq (q - z)\, y$, or

$$py - (q - z)\, y \leq e \tag{27}$$

Left hand side of (27) is Value-at-Risk (the worst possible loss) relative to today's market value of assets, which must be met by equity e.

Since the constraint binds, the optimal holding of the risky securities for the leveraged investor is

$$y = \frac{e}{z - (q - p)} \tag{28}$$

and the balance sheet is

Assets	Liabilities
securities, py	equity, e debt, $(q - z)\,y$

(29)

Aggregation.

Since (28) is linear in e, the aggregate demand of the leveraged sector has the same form as (28) when e is the *aggregate capital* of the leveraged sector as a whole.

Market-clearing

y_A is holding of risky securities by active investors and y_P the holding by the passive investors. Market clearing condition is

$$y_A + y_P = S \tag{30}$$

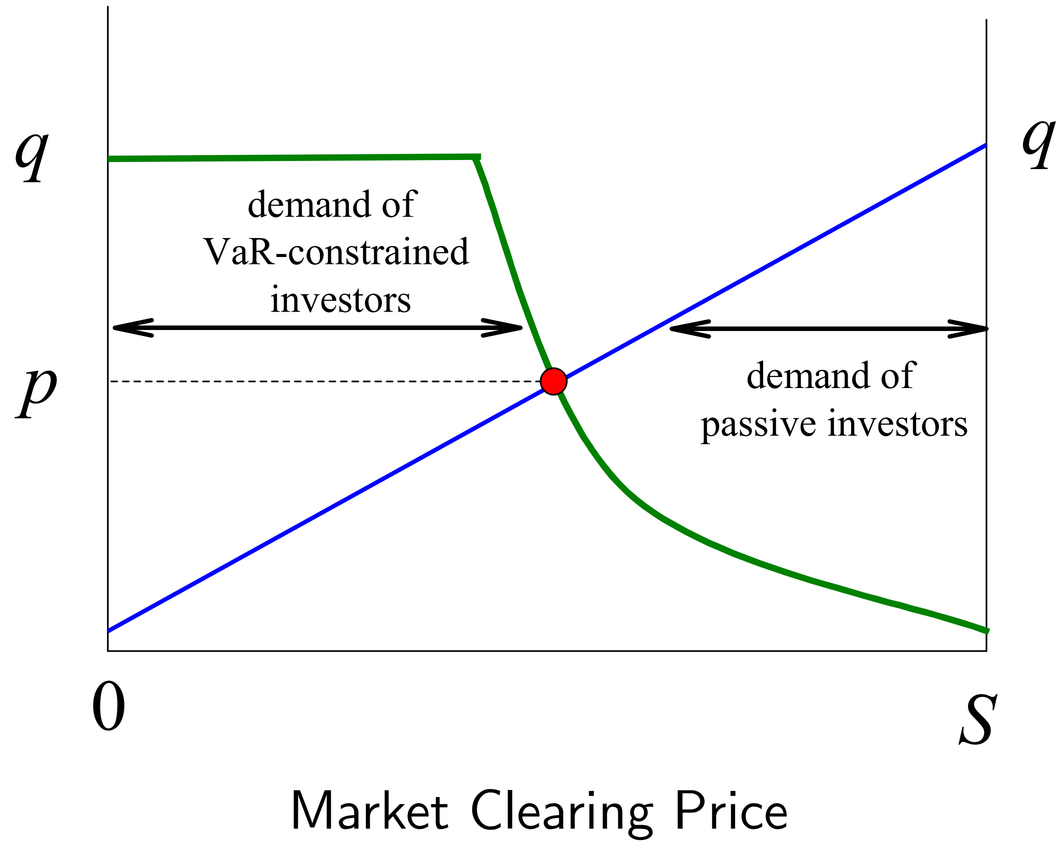
S is the total endowment of the risky securities.

Figure **??** illustrates the equilibrium for a fixed value of aggregate capital e.

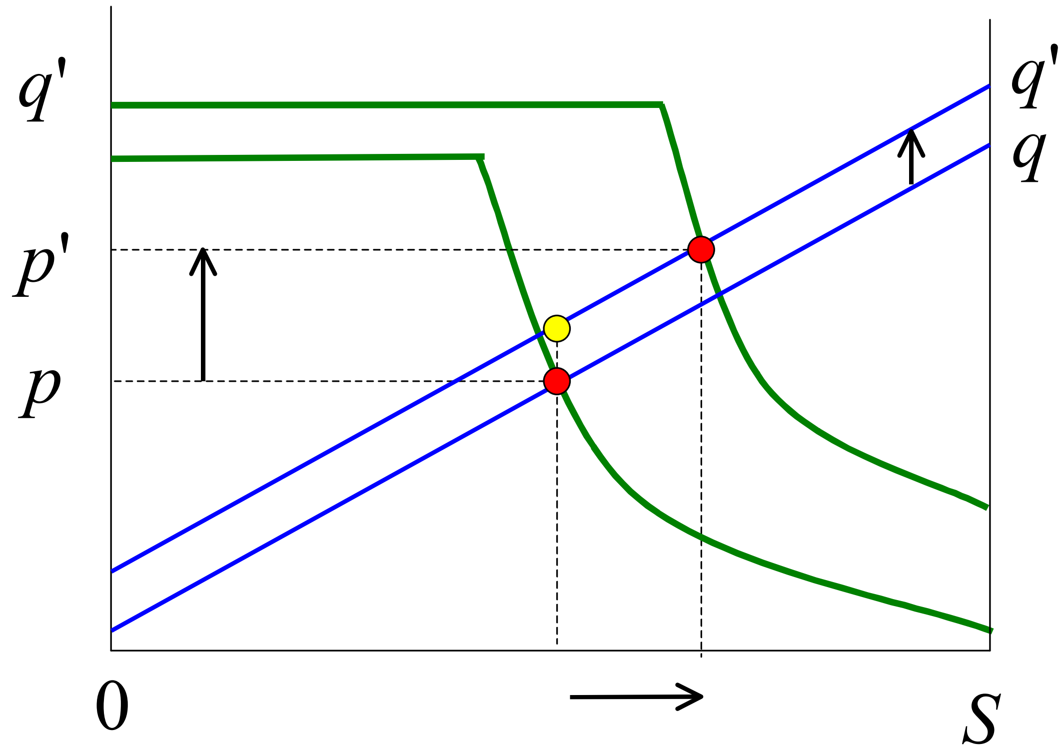
For the passive investors, their demand is linear, with the intercept at q.

The demand of the leveraged sector can be read off from (28).

The solution is fully determined as a function of e. In a dynamic model, e can be treated as the state variable (see Danielsson, et al. (2009)).



q
demand of
VaR-constrained
investors
q
p
demand of
passive investors
0
S
Market Clearing Price



Amplified response to improvement in fundamentals q

Comparative Statics.

Consider an improvement in the fundamentals of the risky security, where expected payoff of the risky securities rises from from q to q'.

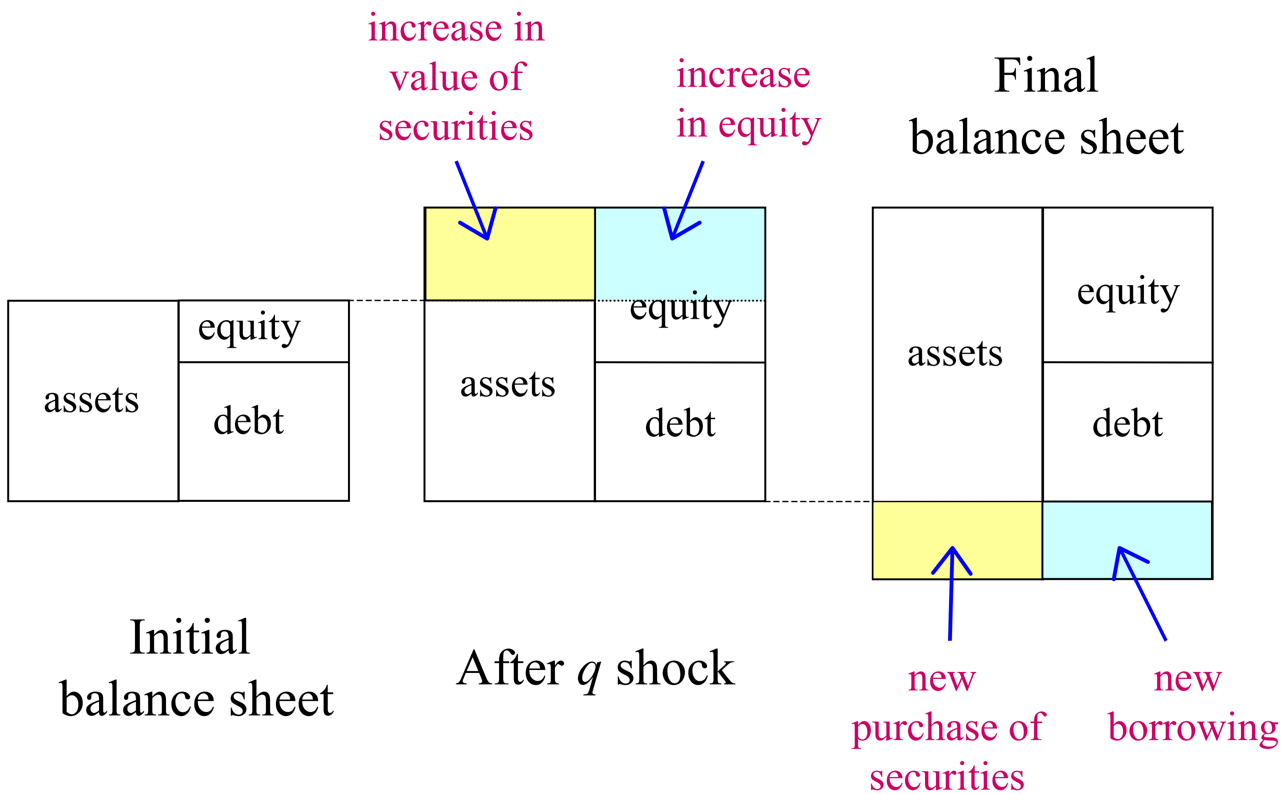
In banking interpretation an improvement in the expected payoff could result from an improvement in the macroeconomic outlook, lowering the probability that the borrowers would default on their loans.

Denote by e' the new equity level of the leveraged investors that incorporates the capital gain when the price rises to p'. The initial amount of debt was $(q-z)\,y$. Since the new asset value is $p'y$, the new equity level e' is

$$\begin{aligned} e' &= p'y - (q-z)\,y \\ &= (z + p' - q)\,y \end{aligned} \tag{31}$$

Initial balance sheet is on the left, where the total asset value is py.

Middle balance sheet shows the effect of an improvement in fundamentals that comes from an increase in q, but before any adjustment in the risky security holding. There is an increase in the value of the securities without any change in the debt value, since the debt was already risk-free to begin with. So, the increase in asset value flows through entirely to an increase in equity. Equation (31) expresses the new value of equity e' in the middle balance sheet in Figure **??**.



Balance sheet expansion from q shock

Increase in equity relaxes the Value-at-Risk constraint, and the leveraged sector can increase its holding of risky securities. The new holding y' is larger, and is enough to make the VaR constraint bind at the higher equity level, with a higher fundamental value q'. That is,

$$\begin{aligned} e' &= p'y' - (q - z)\, y' \\ &= (z + p' - q')\, y' \end{aligned} \tag{32}$$

After the q shock, the investor's balance sheet has strengthened, in that capital has increased without any change in debt value.

There has been an erosion of leverage

There is *spare capacity* on the balance sheet in the sense that equity is now larger than is necessary to meet the Value-at-Risk.

In order to utilize the slack in balance sheet capacity, the investor takes on additional debt to purchase additional risky securities.

The demand response is upward-sloping.

New holding of securities is now y', and the total asset value is $p'y'$.

Equation (32) expresses the new value of equity e' in terms of the new higher holding y' in the right hand side balance sheet in Figure **??**. From (31) and (32), we can write the new holding y' of the risky security as

$$y' = y\left(1 + \frac{q' - q}{z + p' - q'}\right) \tag{33}$$

From the demand of passive investors (25) and market clearing,

$$p' - q' = \frac{z^2}{3\tau}(y' - S)$$

Substituting into (33),

$$y' = y\left(1 + \frac{q' - q}{z + \frac{z^2}{3\tau}\left(y' - S\right)}\right) \tag{34}$$

This defines a quadratic equation in y'. The solution is where the right hand side of (34) cuts the 45 degree line. The leveraged sector amplifies booms and busts if $y' - y$ has the same sign as $q' - q$. Then, any shift in fundamentals gets amplified by the portfolio decisions of the leveraged sector. The condition for amplification is that the denominator in the second term of (34) is positive. But this condition is guaranteed from (33) and the fact that $p' > q' - z$ (i.e. that the price of the risky security is higher than its worst possible realized payoff).

- Note size of the amplification is increasing in leverage, seen from the fact that $y' - y$ is larger when z is small. Recall that z is the fundamental risk. When z is small, the associated Value-at-Risk is also small, allowing the leveraged sector to maintain high leverage. The higher is the leverage, the greater is the marked-to-market capital gains and losses.

- Amplification is large when the leveraged sector itself is large relative to the total economy.

- Finally, note that the amplification is more likely when the passive sector's risk tolerance τ is high.

The price gap, $q - p$ is the difference between the expected payoff from the risky security and its price. It is one measure of the price of risk in the economy. The market clearing condition and the demand of the passive sector (25) give an empirical counterpart to the price gap given by the size of the leveraged sector. Recall that y_A is the holding of the risky security by the leveraged sector. We have

$$q - p = \frac{z^2}{3\tau}(S - y_A) \tag{35}$$

Empirical Hypothesis. Risk premiums are low when the size of the leveraged sector is large relative to the non-leveraged sector.

- Amplifying mechanism works exactly in reverse on the way down.

- A negative shock to the fundamentals of the risky security drives down its price, which erodes the marked-to-market capital of the leveraged sector.

- The erosion of capital induces the sector to shed assets so as to reduce leverage down to a level that is consistent with the VaR constraint.

- Risk premium increases when the leveraged sector suffers losses, since $q - p$ increases.

Pricing of Risk and Credit Supply

For now, treat S (the total endowment of the risky security) as exogenous.

Begin with the market-clearing condition for the risky security, $y_A + y_P = S$. Substituting in the expressions for the demands of the active and passive sectors, we can write the market clearing condition as

$$\frac{e}{z-(q-p)} + \frac{3\tau}{z^2}(q-p) = S \tag{36}$$

We also impose a restriction on the parameters from the requirement that the active investors have a strictly positive total holding of the risky security, or equivalently that the passive sector's holding is strictly smaller than the total endowment S. From (25) this restriction can be written as

$$\frac{3\tau}{z^2}(q-p) < S \tag{37}$$

Our discussion so far of the amplification of shocks resulting from the leveraged investors' balance sheet management suggests that a reasonable hypothesis is that the risk premium to holding the risky security is falling as the fundamental payoff of the risky security improves. This is indeed the case. We have:

Proposition 1. *The expected return on the risky security is strictly decreasing in* q.

The expected return to the risky security is $(q/p) - 1$. It is more convenient to work with a monotonic transformation of the expected return given by

$$\pi \equiv 1 - \frac{p}{q} \tag{38}$$

We see that π lies between zero and one. When $\pi = 0$, the price of the risky security is equal to its expected payoff, so that there is no risk premium in holding the risky security over cash. As π increases, the greater is the expected return to holding the risky security. Using the π notation, the market-clearing condition (36) can be written as follows.

$$F \equiv e + \frac{3\tau}{z^2} q\pi \left(z - q\pi\right) - S\left(z - q\pi\right) = 0 \tag{39}$$

We need to show that π is decreasing in q. From the implicit function theorem,

$$\frac{d\pi}{dq} = -\frac{\partial F/\partial q}{\partial F/\partial \pi} \tag{40}$$

and

$$\frac{\partial F}{\partial q} = \pi \left(\frac{3\tau}{z} \left(1 - \frac{2\pi q}{z} \right) + S \right)$$

Dividing this expression by $3\tau\pi/z^2 > 0$, we see that $\partial F/\partial q$ has the same sign as

$$
\begin{aligned}
& (z - \pi q) + \left(\frac{z^2}{3\tau} S - \pi q \right) \\
= \quad & (z - (q - p)) + \left(\frac{z^2}{3\tau} S - (q - p) \right)
\end{aligned}
\tag{41}
$$

The left hand term in (41) is positive since price p is above the minimum payoff $q - z$. The right hand term is positive from our parameter restriction (37) that ensures that the risky security holding by the leveraged sector is strictly positive. Hence, $\partial F / \partial q > 0$. Similarly, it can be shown that $\partial F / \partial \pi > 0$. Therefore, $d\pi / dq < 0$. This concludes the proof of Proposition 1.

The expected return on the risky security is falling as the fundamentals improve. Risk premium in the economy is declining during booms.

When fundamentals improve, the leveraged investors (the banks) experience mark-to-market gains on their balance sheets, leading to higher equity capital The higher mark-to-market capital generates additional balance sheet capacity for the banks that must be put to use. Excess balance sheet capacity is put to use by increasing lending (purchasing more risky securities) with money borrowed from the passive investors.

Shadow Value of Bank Capital

Lagrange multiplier associated with the constrained optimisation problem of the banks is the rate of increase of the objective function with respect to a relaxation of the constraint, and hence can be interpreted as the shadow value of bank capital. Denoting by λ the Lagrange multiplier, we have

$$\begin{aligned}
\lambda &= \frac{dE(W)}{de} \\
&= \frac{dE(W)}{dy}\frac{dy}{de} \\
&= (q-p) \cdot \frac{1}{z-(q-p)} \qquad (42)
\end{aligned}$$

where we have obtained the expression for $dE(W)/dy$ from (22) and dy/de is obtained from (28), which gives the optimal portfolio decision of the leveraged investor. We see from (42) that as the price gap $q-p$ becomes compressed, the Lagrange multiplier λ declines. The implication is that the marginal increase of a dollar's worth of new capital for the leveraged investor is generating less expected payoff. As the price gap $q-p$ goes to zero, so does the Lagrange multiplier, implying that the return to a dollar's worth of capital goes to zero.

Furthermore, we have from (35) that the price gap $q-p$ is decreasing as the size of the leveraged sector increases relative to the whole economy. The shadow value of bank capital can then be written as

$$\begin{aligned} \lambda &= (q-p) \cdot \frac{1}{z-(q-p)} \\ &= \frac{z\,(S-y_A)}{3\tau + z\,(y_A - S)} \end{aligned} \tag{43}$$

We have the following proposition.

Proposition 2. *The shadow value of bank capital is decreasing in the size of the leveraged sector.*

The *leverage* of the active investor is defined as the ratio of total assets to equity. Leverage is given by

$$\begin{aligned} \frac{py}{e} &= \frac{p}{e} \times \frac{e}{z-(q-p)} \\ &= \frac{p}{z-(q-p)} \end{aligned} \tag{44}$$

As q increases, the numerator $p(q)$ increases without bound. Since the price gap is bounded below by zero, overall leverage eventually increases in q. Thus, leverage is high when total assets are large. In the terminology of Adrian and Shin (2007), the leveraged investors exhibit *pro-cyclical leverage*.

In the run-up to the global financial crisis of 2007 to 2009, the financial system was said to "awash with liquidity", in the sense that credit was easy to obtain.

When asset prices rise, financial intermediaries' balance sheets generally become stronger, and—without adjusting asset holdings—their leverage becomes eroded.

The financial intermediaries then hold **surplus capital**, and they will attempt to find ways in which they can employ their surplus capital.

Analogy with manufacturing firms: financial system as having "surplus capacity". For such surplus capacity to be utilized, the intermediaries must expand their balance sheets. On the liability side, they take on more debt. On the asset side, they search for potential borrowers.

When the set of potential borrowers is fixed, the greater willingness to lend leads to an erosion in risk premium from lending, and spreads become compressed.

Supply of Credit

Now examine endogenous S, and **loan supply.**

Suppose there is a large pool of potential borrowers who wish to borrow to fund a project, from either the active investors (the banks) or the passive investors (the households). They will borrow from whomever is willing to lend.

Assume that the potential borrowers are identical, and each have identical projects to those which are already being financed by the banks and households. In other words, the potential projects that are waiting to be financed are perfect substitutes with the projects already being funded. Denote the risk premium associated with the pool of potential projects by the constant π_0. If the market risk premium were ever to fall below π_0, the investors in the existing projects would be better off selling the existing projects to fund the projects that are sitting on the sidelines. Therefore, the market premium cannot fall below π_0, so that in any equilibrium with endogenous credit supply, we have

$$\pi \geq \pi_0 \tag{45}$$

Define the *supply of credit function* $S(q)$ as the function that maps q to the total lending S. When $\pi(q) \geq \pi_0$, there is no effect of a small change in q on the supply of credit. Define q^* as the threshold value of q defined as $q^* = \pi^{-1}(\pi_0)$. When $q > q^*$, then the equilibrium stock of lending S is determined by the market clearing condition (39) where $\pi = \pi_0$. Hence, S satisfies

$$F \equiv e + \frac{3\tau}{z^2} q\pi_0 (z - q\pi_0) - S(z - q\pi_0) = 0$$

The slope of the supply of credit function is given by

$$\frac{dS}{dq} = -\frac{\partial F/\partial q}{\partial F/\partial S} \tag{46}$$

We know from (41) that the numerator of (46) is positive, while $\partial F/\partial S = -(z - q\pi_0) = q - p - z < 0$. Therefore $dS/dq > 0$, so that credit supply is increasing in q. We can summarise the result as follows.

Proposition 3. *The supply of credit S is strictly increasing in q when $q > \pi^{-1}(\pi_0)$.*

The assumption that the pool of potential borrowers have projects that are perfect substitutes for the existing projects being funded is a strong assumption, and unlikely to hold in practice. Instead, it would be reasonable to suppose that the project quality varies within the pool of potential borrowers, and that the good projects are funded first. For instance, the pool of borrowers could consist of households that do not yet own a house, but would like to buy a house with a mortgage. Among the potential borrowers would be good borrowers with secure and verifiable income.

However, as the good borrowers obtain funding and leave the pool of potential borrowers, the remaining potential borrowers will be less good credits.

If the banks' balance sheets show substantial slack, they will search for borrowers to lend to.

As balance sheets continue to expand, more borrowers will receive funding.

When all the good borrowers already have a mortgage, then the banks must lower their lending standards in order to generate the assets they can put on their balance sheets. In the sub-prime mortgage market in the United States in the years running up to the financial crisis of 2007, we saw that when balance sheets are expanding fast enough, even borrowers that do not have the means to repay are granted credit—so intense is the urge to employ surplus capital. The seeds of the subsequent downturn in the credit cycle are thus sown.