Balance Sheet Capacity and Endogenous Risk*

Jon Danielsson                  Hyun Song Shin
London School of Economics      Princeton University

Jean-Pierre Zigrand
London School of Economics

This version: August 2011

Abstract

Banks operating under Value-at-Risk constraints give rise to a well-defined aggregate balance sheet capacity for the banking sector as a whole that depends on total bank capital. Equilibrium risk and market risk premiums can be solved in closed form as functions of aggregate bank capital. We explore the empirical properties of the model in light of recent experience in the financial crisis and highlight the importance of balance sheet capacity as the driver of the financial cycle and market risk premiums.

JEL codes: G01, G21, G32
Keywords: Banking Crises, Financial Intermediation, Value-at-Risk

*This paper supersedes our earlier paper circulated under the title “Risk Appetite and Endogenous Risk”. We are grateful to Gara Afonso, Rui Albuquerque, Markus Brunnermeier, Hui Chen, Mark Flannery, Antonio Mele, Anna Pavlova, Dimitri Vayanos, Ivo Welch, Wei Xiong and participants at the Adam Smith Asset Pricing Conference, CEMFI, Exeter, Luxembourg, Maastricht, MIT Sloan, the NBER Conference on Quantifying Systemic Risk, Northwestern, the Federal Reserve Bank of New York, Pompeu Fabra, the Fields Institute, Venice and the Vienna Graduate School of Finance for comments on earlier drafts.
1 Introduction

The global financial crisis of 2007-9 has served as a reminder of the important role played by banks and other financial intermediaries in driving the financial cycle. The compression of risk spreads during the upswing of the credit cycle followed by the sharp increase in spreads with the onset of the crisis have been a central theme in the commentary on the crisis. We posit and solve for the notion of the banking sector’s “balance sheet capacity” as the explanatory variable that drives the financial cycle.

In textbook discussions of corporate financing decisions, the set of positive net present value (NPV) projects is often taken as being exogenously given, with the implication that the size of the balance sheet is fixed. The focus is on the liabilities side of the balance sheet, in determining the relative mix of equity and debt in financing the assets. Leverage increases by substituting equity for debt, such as through an equity buyback financed by a debt issue, as depicted by the upper panel in Figure 1.

However, the upper panel in Figure 1 turns out not to be a good description of the way that the banking sector’s leverage varies over the cycle. The distinguishing feature of the banking sector leverage cycle is that leverage fluctuates through fluctuations in the total size of the balance sheet with equity being the forcing variable. Hence, leverage and total assets tend to move in lock-step, as in the lower panel of Figure 1

Balance sheet capacity refers to the capacity of intermediaries to take on risk exposures, which depends on the volatility of asset returns. However, equilibrium volatility is an endogenous variable. It depend on (among other things) the ability of banks to take on risky exposures. So, we have a circularity. Solving for equilibrium entails solving simultaneously for volatility, risk premiums and balance sheet capacity.

We construct a dynamic banking model where the banking sector consists

---

1See Adrian and Shin (2008, 2010) for a discussion of the evidence from the US.
of risk-neutral banks that are subject to a Value-at-Risk (VaR) constraint that requires them to maintain a capital cushion that limits their probability of insolvency at all times to some small constant $\alpha$. We solve for the equilibrium in closed form and examine how balance sheet capacity, volatility and risk premiums are jointly determined.

The Lagrange multiplier associated with the VaR constraint enters into the banks’ lending decisions just like a risk aversion parameter. During booms when banks are well capitalized, banks lower their lending standards to use up spare capacity on their balance sheet. In the downturn, however, as banks suffer erosion of their capital, the capital constraint binds harder so that the Lagrange multiplier associated with their capital constraint increases. Hence, their “as if” risk aversion increases in the downturn. This combination of increasing risk and increased risk aversion leads to a rise in the risk premium in the economy. The expected returns to risky assets increase, and projects that previously received funding from the banking sector

Figure 1: **Two Modes of Leveraging Up.** In the upper panel, the firm keeps assets fixed but replaces equity with debt. In the lower panel, the firm keeps equity fixed and increases the size of its balance sheet.
Figure 2: **Nature of equilibrium.** In equilibrium, there is mutual dependence between balance sheet capacity, volatility and risk premiums.

no longer receives funding.

Our model and its solution have a number of attractive features. First, we are able to solve our dynamic banking model in closed form, where equilibrium volatility and risk premiums can both be solved as functions of a single state variable \( \theta \), where

\[
\theta = \frac{\text{Size of long-only sector}}{\text{Banking sector equity}}
\]

where the numerator is some constant times the holding of the risky asset by the non-bank sector. The denominator is the aggregate capital of the banking sector as a whole. In terms of our state variable \( \theta \), the equilibrium volatility \( \sigma (\theta) \) takes the particularly simple form:

\[
\sigma (\theta) = \left\{ \text{fundamental volatility} \right\} \times \theta \exp \left\{ -\theta \right\} \times F (\theta)
\]

where “fundamental volatility” refers to the volatility due to the exogenous shocks to the economy, and the function \( F (\theta) \) ensures that in the limiting case without a banking sector (i.e. as \( \theta \to \infty \)), the equilibrium volatility converges to the fundamental volatility. Thus, the shape of the volatility...
function $\sigma(\theta)$ is driven by the term $\theta \exp\{-\theta\}$. In particular, we show that there is a cut-off value of $\theta$ above which equilibrium volatility exceeds the fundamental volatility, but below which equilibrium volatility is below fundamental volatility.

In this way, we capture the notion that volatility is high when bank capital is depleted. Black (1976) and Schwert (1989) noted that declining asset prices lead to increased future volatility empirically. Fostel and Geanakoplos (2010, 2011) present a theory for increased volatility in a downturn within a framework that builds on Geanakoplos’s (2009) model of leverage with heterogenous beliefs. To our knowledge, our paper is the first paper to solve for the equilibrium volatility function in closed form in a dynamic banking model.

Second, our framework is flexible enough to extend the analysis to the multi-asset setting. In the multi-asset version of our model, closed form solutions are still available for some special cases. We show that correlations in returns emerge endogenously even though the fundamentals driving the asset returns are independent, and that the correlation can be characterized quite cleanly in terms of the fundamentals.

Indeed, the closed form solution is sufficiently compact that we can address more applied topics such as derivatives pricing and the shape of the volatility curve. Kim and Kon (1994), Tauchen, Zhang and Liu (1996) and Anderson, Bollerslev, Diebold and Ebens (2001) find that the leverage effect is stronger for stock indices than for individual securities. This has been considered as a puzzle for a literal interpretation of the Black (1976) and Schwert (1989) leverage effect. However, in our framework, this feature can be proved as a theorem.

Our paper has implications for policy issues, also. The fact that risk premiums are determined by aggregate banking sector capital is very much in line with recent “macroprudential” thinking among policy makers whose aim is to ensure that banking sector stress tests are in place to ensure that the
banking sector has sufficient capacity to perform its economic role of channeling funding from savers to borrowers. This is in contrast to the previously “microprudential” concern with ensuring that banks have sufficient capital to serve as a buffer against loss that protects depositors (and hence the deposit insurance agency) from losses. Whereas microprudential concerns have to do with avoiding fiscal costs (due to bank recapitalization), macroprudential concerns have to do with maintaining banking sector lending capacity.

The outline of the paper is as follows. We begin with a review of the related literature, and then move to the general statement of the problem. We characterize the closed-form solution of our model for the single risky asset case first. We derive an ordinary differential equation that characterizes the market dynamics in this case, and examine the solution. We then extend the analysis to the general multi-asset case where co-movements can be explicitly studied. We begin with a brief review of the literature in order to place our paper in context.

1.1 Related Literature

Our paper lies at the confluence of two strands in the literature. One strand is the literature on crisis dynamics in competitive equilibrium, such as Gennette and Leland (1990), Geanakoplos (1997, 2009) and Geanakoplos and Zame (2003). The second strand is the corporate finance literature that draws insights on balance sheet constraints, such as Shleifer and Vishny’s (1997) observation that margin constraints limit the ability of arbitrageurs to exploit price differences, as well as Holmström and Tirole’s (1998) work on debt capacities. Our paper draws on both these strands by borrowing tools from corporate finance into the study of asset pricing and financial market fluctuations.

Our paper belongs in the category of recent work where balance sheet constraints enter as a channel of contagion. However, by addressing the up-phase of the financial cycle, and the potential for the compression of
risk premiums during lending booms, our approach differs from the more
common “financial frictions” models where the exogenous leverage constraint
binds only in the downturn. In such models, lending is always below first
best. As well as on the downturn, our focus is on the up-phase of the cycle
when risk premiums become compressed, leaving the economy vulnerable to
a potential reversal.\footnote{Examples of the “financial frictions” literature include Kiyotaki and Moore (1997), Aiyagari and Gertler (1999), Basak and Croitoru (2000), Gromb and Vayanos (2002), Brunnermeier (2008) and Brunnermeier and Pedersen (2009), Garleanu and Pedersen (2009), Chabakauri (2008), Rychkov (2008). Garleanu and Pedersen (2009) extend the CAPM by incorporating a capital constraint to show how assets with the same fundamental risk may trade at different prices. He and Krishnamurthy (2007) study a dynamic asset pricing model with intermediaries, where the intermediaries’ capital constraints enter into the asset pricing problem as a determinant of portfolio capacity. Brunnermeier and Sannikov (2009) share with our paper the focus on financial intermediation although their focus is on examining the effect on the real economy through real investment decisions. Brunnermeier and Pedersen (2009) emphasize the “margin spirals” that result where capital constraints set off amplified feedback effects.}

Amplification through wealth effects was studied by Xiong (2001), Kyle
and Xiong (2001) who show that shocks to arbitrageur wealth can amplify
volatility when the arbitrageurs react to price changes by rebalancing their
portfolios. In a multi-asset and multi-country centre-periphery extension,
Pavlova et al (2008) find that wealth effects across countries are strengthened
further if the center economy faces portfolio constraints. In these papers,
margin constraints are time-varying and can serve to amplify market fluctua-
tions through reduced risk-bearing capacity, and therefore behave more like
the risk-sensitive constraints we study below. Incorporating balance sheet
constraints on asset pricing problems have been examined by Adrian, Etula
and Shin (2009) for the foreign exchange market, Etula (2009) for the com-
modities market and by Adrian, Moench and Shin (2009) for the interaction
between macro and balance sheet variables.

Most directly, this paper builds on the work on Lagrange multiplier asso-
ciated with Value-at-risk constraints. Our earlier paper (Danielsson, Shin
and Zigrand (2004)) had backward-looking learning rather than solving for
equilibrium in a rational expectations model. Brunnermeier and Pedersen (2009), Oehmke (2008) and Garleanu and Pedersen (2009) have explored the consequences of fluctuating Lagrange multipliers associated with balance sheet constraints. There is a small but growing empirical literature on risk appetite. Surveys can be found in Deutsche Bundesbank (2005) and in BIS (2005, p. 108). See also Coudert et al. (2008) who argue that risk tolerance indices (such as the Global Risk Aversion Index (GRAI), the synthetic indicator LCVI constructed by J.P. Morgan, PCA etc) tend to predict stock market crises. Gai and Vause (2005) provide an empirical method that can help distinguish risk appetite from the related notions of risk aversion and the risk premium.

Relative to the earlier papers, our incremental contribution is to solve for the equilibrium returns, volatility and correlations in closed form as a fixed point of the equilibrium mapping. Furthermore, the closed-form solution takes a particularly simple form, depending on the size of the banking sector relative to the financial system as a whole. The tractability afforded by our closed-form solution is instrumental in deriving several of the insights in our paper, and opens up a number of useful avenues to link the banking literature with insights from asset pricing.

2 The Model

Our model describes the interactions between two groups of investors - passive investors and active investors. The passive investors can be thought of as value investors such as households, pension funds and mutual funds, while the active investors can be interpreted as banks and other intermediaries.

The risky securities can be interpreted as loans granted to ultimate borrowers, but where there is a risk that the borrowers do not fully repay the loan. Figure 3 depicts the relationships. Under this interpretation, the market value of the risky securities can be thought of as the marked-to-market
value of loans granted to the ultimate borrowers. The value investors’ holding of the risky security can be interpreted as the credit that is granted directly by the household sector (through the holding of corporate bonds, for example), while the holding of the risky securities by the active investors can be given the interpretation of intermediated credit through the banking sector.

Let time be indexed by \( t \in [0, \infty) \). There are \( N > 0 \) non-dividend paying risky assets as well as a risk-free bond. We will focus later on the case where \( N = 1 \), but we state the problem for the general \( N \) asset case. The price of the \( i \)-th risky asset at date \( t \) is denoted \( P^i_t \). We will look for an equilibrium in which the price processes for the risky assets follow:

\[
\frac{dP^i_t}{P^i_t} = \mu^i_t dt + \sigma^i_t dW_t ; \quad i = 1, \ldots, N
\]  

(3)

where \( W_t \) is an \( N \times 1 \) vector of independent Brownian motions, and where the scalar \( \mu^i_t \) and the \( 1 \times N \) vector \( \sigma^i_t \) are as yet undetermined processes that will be solved in equilibrium. The risk-free bond has price \( B_t \) at date \( t \), which is given by \( B_0 = 1 \) and \( dB_t = rB_t dt \), where \( r \) is constant.
2.1 Portfolio Choice of Banks

The banks (the financial intermediaries, or “FIs”) have short horizons who maximize the instantaneous expected returns on their loan portfolio. But each bank is subject to a risk constraint where its capital $V$ is sufficiently large to cover its Value-at-Risk (VaR). We use “capital” and “equity” interchangeably in what follows.

We do not provide further microfoundations for the VaR rule here, but capital budgeting practices based on measured risks (such as VaR) are well-established among banks, and we adopt it here as a key feature of our model. The short-horizon nature of our model is admittedly stark, but can be seen as reflecting the same types of frictions that give rise to the use of constraints such as VaR, and other commonly observed institutional features among banks and other large financial institutions. Finally, note that we have denoted the bank’s capital as $V$ without a subscript for the bank, as it will turn out that there is a natural aggregate result where only the aggregate banking sector capital matters for equilibrium, rather than the distribution of bank capital.

Let $a_t^i$ be the number of units of the $i$th risky asset held at date $t$, and denote the dollar amount invested in risky security $i$ by

$$D_t^i := a_t^i P_t^i$$  \hspace{1cm} (4)

The budget constraint of the trader is

$$b_t B_t = V_t - a_t^T P_t = V_t - \sum_{i=1}^{N} D_t^i$$  \hspace{1cm} (5)

where $V_t$ is the trader’s capital and where $x^T$ is the transpose of $x$. The

---

3See Adrian and Shin (2008) for one possible microfoundation in a contracting model with moral hazard, and Danielsson and Zigrand (2008) for a forward looking general equilibrium model with production where a VaR constraint reduces the probability of a systemic event caused by a free-riding externality during the refinancing stage.
dynamic budget constraint governs the evolution of capital in the usual way:

\[
dV_t = a_t^\top dP_t + b_t dB_t = [rV_t + D_t^\top (\mu_t - r)] dt + D_t^\top \sigma_t dW_t
\]

where \( D_t^\top \) denotes the transpose of \( D_t \), and where \( \sigma_t \) is the \( N \times N \) diffusion matrix, row \( i \) of which is \( \sigma_i^t \). In (6), we have abused notation slightly by writing \( r = (r, \ldots, r) \) in order to reduce notational clutter. The context should make it clear where \( r \) is the scalar or the vector.

From (6), the expected capital gain is

\[
E_t[dV_t] = [rV_t + D_t^\top (\mu_t - r)] dt
\]

and the variance of the trader’s equity is

\[
\text{Var}_t(dV_t) = D_t^\top \sigma_t \sigma_t^\top D_t dt
\]

We assume (and later verify in equilibrium) that the variance-covariance matrix of instantaneous returns is of full rank and denote it by

\[
\Sigma_t := \sigma_t \sigma_t^\top
\]

The bank is risk-neutral, and maximizes return (7) subject to its VaR constraint, where VaR is some constant \( \alpha \) times the forward-looking standard deviation of returns on equity. It is without loss of generality to define Value-at-Risk in this way due to the Gaussian nature of the Brownian shocks. The number \( \alpha \) is just a normalizing constant, and does not enter materially into the analysis.

Motivated by the evidence from the scatter chart of asset growth and leverage growth of the Wall Street investment banks, we take the bank’s capital \( V_t \) as the state variable. Assuming that the bank is solvent (i.e. \( V_t > 0 \)), the maximization problem can be written as:

\[
\max_{D_t} rV_t + D_t^\top (\mu_t - r) \quad \text{subject to} \quad \alpha \sqrt{D_t^\top \sigma_t \sigma_t^\top D_t} \leq V_t
\]
Once the dollar values \( \{D^i_t\}_{i=1}^N \) of the risky assets are determined, the bank’s residual bond holding is determined by the balance sheet identity:

\[
b_t B_t = V_t - \sum_i D^i_t
\]  

(11)

The first-order condition for the optimal \( D \) is

\[
\mu_t - r = \alpha(D^\top_t \Sigma_t D_t)^{-1/2} \gamma_t \Sigma_t D_t
\]  

(12)

where \( \gamma_t \) is the Lagrange multiplier associated with the VaR constraint. Hence,

\[
D_t = \frac{1}{\alpha(D^\top_t \Sigma_t D_t)^{-1/2} \gamma_t} \Sigma_t^{-1}(\mu_t - r)
\]  

(13)

When \( \mu_t \neq r \), as will occur in equilibrium, the objective function is monotonic in \( D_t \) by risk-neutrality, and the constraint must bind. Hence,

\[
V_t = \alpha \sqrt{D^\top_t \Sigma_t D_t}
\]  

(14)

and therefore

\[
D_t = \frac{V_t}{\alpha^2 \gamma_t} \Sigma_t^{-1}(\mu_t - r)
\]  

(15)

Notice that the optimal portfolio is similar to the mean-variance optimal portfolio allocation, where the Lagrange multiplier \( \gamma_t \) appears in the denominator, just like a risk-aversion coefficient. We thus have a foretaste of the main theme of the paper - namely, that the banks in our model are risk-neutral, but they will behave like risk-averse investors whose risk aversion appears to shift in line with the Lagrange multiplier \( \gamma \). Substituting into (14) and rearranging we have

\[
\gamma_t = \frac{\sqrt{\xi_t}}{\alpha}
\]  

(16)

where

\[
\xi_t := (\mu_t - r)^\top \Sigma_t^{-1} \gamma_t (\mu_t - r) \geq 0
\]  

(17)
The Lagrange multiplier $\gamma_t$ for the VaR constraint is thus proportional to the generalized Sharpe ratio $\sqrt{\xi}$ for the risky assets in the economy. Although traders are risk-neutral, the VaR constraint makes them act as if they were risk-averse with a coefficient of relative risk-aversion of $\alpha^2 \gamma_t = \alpha \sqrt{\xi_t}$. As $\alpha$ becomes small, the VaR constraint binds less and banks’ willingness to take on risk increases.

Notice that the Lagrange multiplier $\gamma_t$ does not depend directly on equity $V_t$. Intuitively, an additional unit of capital relaxes the VaR constraint by a multiple $\alpha$ of standard deviation, leading to an increase in the expected return equal to a multiple $\alpha$ of the generalized Sharpe ratio, i.e. the risk-premium on the portfolio per unit of standard deviation. This should not depend on $V_t$ directly, and indeed we can verify this fact from (17).

Finally, we can solve for the risky asset holdings as

$$D_t = \frac{V_t}{\alpha \sqrt{\xi_t}} \Sigma_t^{-1} (\mu_t - r)$$

(18)

The optimal holding of risky assets is homogeneous of degree one in equity $V_t$. This simplifies our analysis greatly, and allows us to solve for a closed form solution for the equilibrium. Also, the fact that the Lagrange multiplier depends only on market-wide features and not on individual capital levels simplifies our task of aggregation across traders and allows us to view demand (18) without loss of generality as the aggregate demand by the FI sector with aggregate capital of $V_t$.

### 2.2 Closing the Model with Value Investors

We close the model by introducing value investors who supply downward-sloping demand curves for the risky assets. The slope of the value investors’ demand curves will determine the size of the price feedback effect. Suppose that the value investors in aggregate have the following vector-valued
exogenous demand schedule for the risky assets, \( y_t = (y_t^1, \ldots, y_t^N) \) where

\[
y_t = \Sigma_t^{-1} \begin{bmatrix}
\delta^1 (z_t^1 - \ln P_t^1) \\
\vdots \\
\delta^N (z_t^N - \ln P_t^N)
\end{bmatrix}
\]

(19)

where \( P_t^i \) is the market price for risky asset \( i \) and where \( dz_t^i \) is a (favorable) Itô demand shock to the demand of asset \( i \) (or a unfavorable supply shock to security \( i \)) to be specified further. Each demand curve can be viewed as a downward sloping demand hit by demand shocks, with \( \delta^i \) being a scaling parameter that determines the slope of the demand curve. The particular form adopted for these exogenous demands is to aid tractability of the equilibrium pricing function, as we will see shortly. We can interpret these demands as coming from value investors who wish to hold a portfolio of the risky securities where their holding depends on the expected upside return, \( \ln(P_t^{*i}/P_t^i) \), relative to benchmark prices \( P_t^{*i} \) which are given by \( e^{z_t^i} \). The coefficients \( \delta \) play the role of risk tolerance parameters.

Bringing together the demands of the banks and the value investors, the market-clearing condition \( D_t + y_t = 0 \) can be written as

\[
\frac{V_t}{\alpha \sqrt{\xi_t}} (\mu_t - r) + \begin{bmatrix}
\delta^1 (z_t^1 - \ln P_t^1) \\
\vdots \\
\delta^N (z_t^N - \ln P_t^N)
\end{bmatrix} = 0
\]

(20)

Equilibrium prices are therefore

\[
P_t^i = \exp \left( \frac{V_t}{\alpha \delta^i \sqrt{\xi_t}} (\mu_t^i - r) + z_t^i \right); \quad i = 1, \ldots, N
\]

(21)

In solving for the rational expectations equilibrium (REE) of our model, our strategy is to begin with some exogenous stochastic process that drives the passive traders’ demands for the risky assets (the fundamental “seeds” of the model, so to speak), and then solve for the endogenously generated stochastic process that governs the prices of the risky assets.
In particular, we will look for an equilibrium in which the price processes for the risky assets are of the form:

\[ \frac{dP_i^t}{P_i^t} = \mu_i^tdt + \sigma_i^tdW_t \quad ; i = 1, \ldots, N \]  \hspace{1cm} (22)

where \( W_t \) is an \( N \times 1 \) vector of independent Brownian motions, and where the scalar \( \mu_i^t \) and \( 1 \times N \) vector \( \sigma_i^t \) are as yet undetermined coefficients that will be solved in equilibrium. The “seeds” of uncertainty in the equilibrium model are given by the demand shocks of the value investors:

\[ dz_i^t = r^*dt + \eta \sigma_z^t dW_t \]  \hspace{1cm} (23)

where \( \sigma_z^t \) is a \( 1 \times N \) vector that governs which Brownian shocks will get imputed into the demand shocks and therefore govern the correlation structure of the demand shocks. We assume that the stacked \( N \times N \) matrix \( \sigma_z \) is of full rank and that \( r^* > r \), so that demand shocks reflect risk aversion of the value investors.

Our focus is on the way that the (endogenous) diffusion terms \( \{ \sigma_i^t \} \) of the return process depends on the (exogenous) shock terms \( \{ \sigma_z^t \} \), and how the exogenous noise terms may be amplified in equilibrium via the risk constraints of the active traders. Indeed, we will see that the relationship between the two sets of diffusions generate a rich set of empirical predictions.

### 3 Equilibrium with Single Risky Asset

Before examining the general problem with \( N \) risky assets, we first solve the case of with single risky asset. We will look for an equilibrium where the price of the risky asset follows the process:

\[ \frac{dP_t}{P_t} = \mu_t dt + \sigma_t dW_t \]  \hspace{1cm} (24)

where \( \mu_t \) and \( \sigma_t \) are, as yet, undetermined coefficients to be solved in equilibrium, and \( W_t \) is a standard scalar Brownian motion. The “seeds” of
uncertainty in the model are given by the exogenous demand shocks to the value investors’ demands:

\[ dz_t = r^* dt + \eta \sigma_z dW_t \]  

(25)

where \( \sigma_z > 0 \) and \( \eta > 0 \) are known constants. For the single risky asset case, note that

\[ \xi_t = \frac{(\mu_t - r)^2}{\sigma_t^2} \]  

(26)

Substituting into (21), and confining our attention to regions where the Sharpe ratio \( \frac{\mu_t - r}{\sigma_t} \) is strictly positive, we can write the price of the risky asset as

\[ P_t = \exp \left( z_t + \frac{\sigma_t V_t}{\alpha \delta} \right) \]  

(27)

From (24) we have, by hypothesis,

\[ d \ln P_t = \left( \mu_t - \frac{1}{2} \sigma_t^2 \right) dt + \sigma_t dW_t \]  

(28)

Meanwhile, taking the log of (27) and applying Itô’s Lemma gives

\[ d \ln P_t = d \left( z_t + \frac{\sigma_t V_t}{\alpha \delta} \right) \]

\[ = r^* dt + \eta \sigma_z dW_t + \frac{1}{\alpha \delta} d(\sigma_t V_t) \]

\[ = r^* dt + \eta \sigma_z dW_t + \frac{1}{\alpha \delta} \left( \sigma_t dV_t + V_t d\sigma_t + dV_t d\sigma_t \right) \]  

(29)

Now use Itô’s Lemma on \( \sigma(V_t) \):

\[ d\sigma_t = \frac{\partial \sigma}{\partial V_t} dV_t + \frac{1}{2} \frac{\partial^2 \sigma}{\partial (V_t)^2} (dV_t)^2 \]

\[ = \left\{ \frac{\partial \sigma}{\partial V_t} \left[ rV_t + \frac{V_t (\mu_t - r)}{\alpha \sigma_t} \right] + \frac{1}{2} \frac{\partial^2 \sigma}{\partial (V_t)^2} \left( \frac{V_t}{\alpha} \right)^2 \right\} dt + \frac{\partial \sigma}{\partial V_t} \frac{V_t}{\alpha} dW_t \]  

(30)
where (30) follows from

\[
dV_t = [rV_t + D_t(\mu_t - r)]dt + D_t\sigma_t dW_t
\]

\[
= \left[ rV_t + \frac{V_t(\mu_t - r)}{\alpha\sigma_t} \right] dt + \frac{V_t}{\alpha} dW_t
\]  

(31)

and the fact that \( D_t = \frac{V_t}{\alpha\sigma_t} \) due to the binding VaR constraint. Notice that \((dV_t)^2 = \left( \frac{V_t}{\alpha} \right)^2 dt \). We thus obtain diffusion equations for \( V_t \) and for \( \sigma_t \) itself.

Substituting back into (29) and regrouping all \( dt \) terms into a new drift term:

\[
d\ln P_t = \text{(drift term)} \ dt + \left[ \eta\sigma_z + \frac{1}{\alpha\delta} \left( \frac{\sigma_t V_t}{\alpha} + V_t \frac{\partial\sigma_t V_t}{\partial V_t} \right) \right] dW_t
\]  

(32)

We can solve for the equilibrium diffusion \( \sigma_t \) by comparing coefficients between (32) and (28). We have an equation for the equilibrium diffusion given by:

\[
\sigma(V_t) = \eta\sigma_z + \frac{1}{\alpha\delta} \left( \frac{\sigma_t V_t}{\alpha} + V_t \frac{\partial\sigma_t V_t}{\partial V_t} \right)
\]  

(33)

which can be written as the ordinary differential equation (ODE):

\[
V_t^2 \frac{\partial\sigma}{\partial V_t} = \alpha^2 \delta(\sigma_t - \eta\sigma_z) - V_t \sigma_t
\]  

(34)

It can be verified by differentiation that the generic solution to this ODE is given by

\[
\sigma(V_t) = \frac{1}{V_t} e^{\frac{-\sigma}{\sqrt{V_t}}} \left[ c - \alpha^2 \delta \eta\sigma_z \int_{\frac{-\sigma}{\sqrt{V_t}}}^{\infty} e^{-u} du \right]
\]  

(35)

where \( c \) is a constant of integration.

We can set \( c = 0 \) through the following natural restriction in our model. The only randomness in our economy stems from the shocks to the value investor demands. If we let \( \delta \to 0 \), value investors’ demand goes to zero and we get the REE \( \sigma(V_t) = \frac{c}{V_t} \). Since the limit economy should be riskless, we require that returns also are riskless, \( \sigma(V_t) = 0 \), implying that \( c = 0 \).
We thus obtain a unique closed form solution to the rational expectations equilibrium for the single risky asset case. Setting \( c = 0 \) and simplifying, we arrive at the following succinct closed form solution

\[
\sigma(V_t) = \eta \sigma_z \frac{\alpha^2 \delta}{V_t} \exp \left\{ -\frac{\alpha^2 \delta}{V_t} \right\} \times \text{Ei} \left( \frac{\alpha^2 \delta}{V_t} \right) \tag{36}
\]

where \( \text{Ei} (w) \) is the well-known\(^4\) exponential integral function:

\[
\text{Ei} (w) \equiv - \int_{-w}^{\infty} \frac{e^{-u}}{u} du \tag{37}
\]

The \( \text{Ei} (w) \) function is defined provided \( w \neq 0 \). The expression \( \alpha^2 \delta / V_t \) which appears prominently in the closed form solution (36) can be interpreted as the relative scale or size of the value investor sector (parameter \( \delta \)) compared to the banking sector (total capital \( V_t \) normalized by VaR).

To bring out the structure of the solution better, let \( \theta \) be the relative size of the long-only sector relative to the banking sector by defining \( \theta \) as

\[
\theta_t \equiv \frac{\alpha^2 \delta}{V_t} \tag{38}
\]

Then, the closed form solution for equilibrium volatility can be written as a function of \( \theta_t \) as:

\[
\sigma(\theta_t) = \eta \sigma_z \theta_t \exp \{ -\theta_t \} \times \text{Ei} (\theta_t) \tag{39}
\]

The closed form solution also reveals much about the basic shape of the volatility function \( \sigma(\theta_t) \). Consider the limiting case when the banking sector is very small, that is, \( \theta_t \to \infty \). Although \( \theta_t \) becomes large, the exponential term \( \exp \{ -\theta_t \} \) in (39) goes to zero much faster, and so the product \( \theta_t \exp \{ -\theta_t \} \) becomes small. However, since we have exogenous shocks to the value investor demands, there should still be non-zero volatility at the limit, given by the fundamental volatility \( \eta \sigma_z \). The role of the exponential integral term \( \text{Ei} (\theta_t) \) is to ensure that the limiting volatility when the banking

\(^4\)See http://mathworld.wolfram.com/ExponentialIntegral.html
sector becomes small is given by the volatility that would hold in the absence of a banking sector.

The equilibrium risk premium in our model is given by the drift $\mu_t$ (the expected instantaneous return on the risky asset) which can be solved in closed form, and is given by

$$\mu_t = r + \frac{\sigma_t}{2\alpha \eta \sigma_z} \left\{ 2\alpha (r^* - r) + \alpha \sigma_t^2 - \eta \sigma_z + (\sigma_t - \eta \sigma_z) \left[ 2\alpha^2 r + \frac{\alpha^2 \delta}{V_t} - 2 \right] \right\}$$

(40)

We can see that $\mu_t$ depends on the diffusion $\sigma_t$, so that when the expression in the square brackets is positive, $\mu_t$ is increasing in $\sigma_t$. Thus, even though banks are risk-neutral, they are prevented by their VaR constraint from fully exploiting all positive expected return opportunities. The larger is $\sigma_t$, the tighter is the risk constraint, and hence the higher is the expected return $\mu_t$. Note that the expression in the square brackets is positive when $V_t$ is small, which is consistent with the VaR constraint binding more tightly.

Also, notice that as the VaR constraint becomes tighter, $\lim_{\alpha \to \infty} \sigma_t = \eta \sigma_z$ and $\lim_{\alpha \to \infty} \alpha (\sigma_t - \eta \sigma_z) = 0$ so that in the limit we have $\mu_t - \left( \frac{\eta \sigma_z}{2} \right) = r^*$, confirming our interpretation of $r^*$ as the value investor sector’s benchmark log-return.

The information contained in the risk premium $\mu_t$ and its relationship with the diffusion $\sigma_t$ can be summarized alternatively in terms of the Sharpe ratio, which can be written as

$$\frac{\mu_t - r}{\sigma_t} = \frac{1}{2\alpha \eta \sigma_z} \left\{ 2\alpha (r^* - r) + \alpha \sigma_t^2 - \eta \sigma_z + (\sigma_t - \eta \sigma_z) \left[ 2\alpha^2 r + \frac{\alpha^2 \delta}{V_t} - 2 \right] \right\}$$

(41)

The countercyclical shape of the Sharpe Ratio follows directly from the shape of the diffusion coefficient $\sigma_t$. 

19
3.1 Numerical Example

We illustrate the properties of our closed form solution by means of a numerical example. Figure 4 plots the equilibrium diffusion $\sigma_t$ and the drift $\mu_t$ as a function of the state variable $V_t$. The parameters chosen for this plot were $r = 0.01$, $r^* = 0.047$, $\delta = 6$, $\alpha = 2.7$, $\sigma_z = 0.3$, $\eta = 1.5$

As suggested by the closed form solution (36), the plot of $\sigma_t$ is non-monotonic, with a peak when $V_t$ is low. Also, note that when $V = 0$, we have $\sigma = 0.45$, which is the fundamental volatility given by the product of $\sigma_z$ and $\eta$ ($= 0.3 \times 1.5$). The basic non-monotonic shape of the volatility function is quite general, and does not depend on the parameters chosen. We provide further arguments in the appendix.\(^5\)

What Figure 4 reveals is that the feedback effect generating endogenous volatility is strongest for an intermediate value of $V_t$. This is so, since there are two countervailing effects. If $V_t$ is very small - close to zero, say - then there is very little impact of the banks’ portfolio decision on the price of the security. Therefore, both $\sigma_t$ and $\mu_t$ are small. At the opposite extreme, if $V_t$

\(^5\)See Mele (2007) for a discussion of the stylized facts, and for a model generating countercyclical statistics in a more standard framework.
is very large, then banks begin to act more and more like an unconstrained trader. Since the trader is risk–neutral, the expected drift $\mu_t$ is pushed down to the risk–free rate, and the volatility $\sigma_t$ declines.

However, at an intermediate level of $V_t$, the feedback effect is maximized, where a positive price shock leads to greater purchases, which raises prices further, which leads to greater purchases, and so on. This feedback effect increases the equilibrium volatility $\sigma_t$. Due to the risk constraint, the risk-neutral banks behave “as if” they were risk averse, and the equilibrium drift $\mu_t$ reflects this feature of the model. The risk premium $\mu_t$ rises with $\sigma_t$, since both risk and risk aversion increase as bank equity is depleted.

Indeed, as we have commented already, the Lagrange multiplier associated with the risk constraint is the Sharpe ratio in this simple one asset context. The Lagrangian is plotted in Figure 5. We see that the Sharpe ratio rises and falls roughly the same pattern with $\sigma_t$ and $\mu_t$. However, the notable feature of Figure 5 is that the Lagrange multiplier may actually start increasing again when $V$ is large. This is because the Lagrange multiplier reflects the bank’s return on equity (ROE), and ROE is affected by the degree of leverage taken on by the bank. When $V$ becomes large, the volatility falls so that bank leverage increases. What Figure 5 shows is that the increased leverage may
start to come into play for large values of $V$.

Figure 6 gives scatter charts for the relationship between the asset growth and leverage for four sample realizations of the model. Each scatter chart for a particular sample path is accompanied by the price series for that sample path.

The scatter charts reveal the characteristic clustering of dots around the 45-degree line, as shown by Adrian and Shin (2010) for the Wall Street investment banks. The notable feature from the scatter charts is how the slope and degree of clustering depends on the price realizations. When the price path is low, many of the observations are for the upward-sloping part of the volatility function $\sigma(V)$. Along the upward-sloping part, equity depletion associated with price declines is accompanied by a decline in Value-at-Risk, and hence an uptick in leverage. These observations are those below the 45-degree line, but where leverage goes up. However, when the realizations are mainly those on the downward-sloping part of the volatility curve, the scatter chart hugs more closely the 45-degree line. The second example in Figure 6 shows this best.

4 Equilibrium with Many Risky Assets

4.1 General Specification

We now turn to the case with $N > 1$ risky assets and look for an equilibrium in which the prices of risky assets follow:

$$\frac{dP^i_t}{P^i_t} = \mu^i_t dt + \sigma^i_t dW_t$$

(42)

where $W_t$ is an $N \times 1$ vector of independent Brownian motions, and where $\mu^i_t$ and $\sigma^i_t$ are terms to be solved in equilibrium. The demand shocks of the passive traders are given by

$$dz^i_t = r^i dt + \eta^i dW_t$$

(43)
Figure 6: Scatter Charts of Asset and Leverage Changes
where $\sigma^i_t$ is a $1 \times N$ vector that governs which Brownian shocks affect the passive traders’ demands.

We denote conjectured quantities with a tilde. For instance, conjectured drift and diffusion terms are $\tilde{\mu}, \tilde{\sigma}$ respectively and the actual drift and diffusions are $\mu$ and $\sigma$ respectively. For notational convenience, we define the scaled reward-to-risk factor

$$\lambda_t := \frac{1}{\sqrt{\xi_t}} \Sigma_t^{-1} (\mu_t - r)$$

(44)

Also, we use the following shorthands:

$$\beta^i_t := \frac{1}{\sqrt{\xi_t}} (\mu^i_t - r)$$

(45)

$$\gamma^i_t := \frac{1}{\alpha^2 \delta^i} \beta^i_t + \frac{V_t}{\alpha^2 \delta^i} \partial \beta^i_t$$

(46)

and where

$$\frac{\partial \gamma^i_t}{\partial V_t} = \frac{2}{\alpha^2 \delta^i} \partial \beta^i_t + \frac{V_t}{\alpha^2 \delta^i} \partial^2 \beta^i_t$$

Under some conditions to be verified, we can compute the actual drift and diffusion terms of $dP_t^i / P_t^i$ as a function of the conjectured drift and diffusion terms. By Itô’s Lemma applied to (21) we have:

$$\sigma^i_t = \tilde{\sigma}^T V_t \lambda_t \tilde{\sigma}_t + \eta \sigma^i_z$$

(47)

We denote the $N \times 1$ vector of ones by $1_N$, and the operator that replaces the main diagonal of the identity matrix by the vector $v$ by $\text{Diag}(v)$. Also, for simplicity we write $r$ for $r1_N$. Then we can stack the drifts into the vector $\mu_t$, the diffusion coefficients into a matrix $\sigma_t$, etc.

We can solve the fixed point problem by specifying a beliefs updating process $(\tilde{\mu}_t, \tilde{\sigma}_t)$ that when entered into the right hand side of the equation, generates the true return dynamics. In other words, we solve the fixed point problem by solving for self-fulfilling beliefs $(\tilde{\mu}_t, \tilde{\sigma}_t)$ in the equation:

$$\begin{bmatrix} \tilde{\mu}_t \\ \tilde{\sigma}_t \end{bmatrix} = \begin{bmatrix} \mu_t(\tilde{\mu}_t, \tilde{\sigma}_t) \\ \sigma_t(\tilde{\mu}_t, \tilde{\sigma}_t) \end{bmatrix}.$$
By stacking into a diffusion matrix, at a REE the diffusion matrix satisfies

\[ \sigma_t = V_t \epsilon_t \lambda_t^T \sigma_t + \eta \sigma_z \]  

(49)

Using the fact that \( \lambda_t^T \sigma_t \sigma_t^T = \beta_t^T \), \( \sigma_t \) satisfies the following matrix quadratic equation

\[ \sigma_t \sigma_t^T = \eta \sigma_z \sigma_t^T + V_t \epsilon_t \beta_t^T \]  

so that

\[ (\sigma_t - \eta \sigma_z) \sigma_t^T = V_t \epsilon_t \beta_t^T \]  

(50)

The return diffusion in equilibrium is equal to the fundamental diffusion \( \eta \sigma_z \) – the one occurring with no active FIs in the market – perturbed by an additional low-rank term that incorporates the rational equilibrium effects of the FIs on prices. Therefore, we have a decomposition of the diffusion matrix into that part which is due to the fundamentals of the economy, and the part which is due to the endogenous amplification that results from the actions of the active traders. The decomposition stems from relation (47) (keeping in mind that \( \frac{V_t}{\alpha} \lambda_t \sigma_t \) equals the diffusion term of equity)

\[
\sigma_t^i = \left( \frac{1}{\alpha \delta^i} \beta_t^i \right) \text{(vol of capital)} + \left( \frac{V_t}{\alpha \delta^i} \frac{\partial \beta_t^i}{\partial V_t} \right) \text{(vol of capital)} + \eta \sigma_z^i
\]

feedback effect on vol from VaR

feedback effect on vol from changing expectations

We now solve for a representation of \( \sigma_t \). Solutions to quadratic matrix equations can rarely be guaranteed to exist, much less being guaranteed to be computable in closed form. We provide a representation of the solution, should a solution exist. This solution diffusion matrix can be shown to be nonsingular, guaranteeing endogenously complete markets by the second fundamental theorem of asset pricing.

Denote the scalar

\[ e_t := 1 - V_t \lambda_t^T \epsilon_t \]

It follows from the Sherman-Morrison theorem (Sherman and Morrison (1949)) that \( e_t = \text{Det}[I - V_t \epsilon \lambda_t^T] \) and that if (and only if) \( e_t \neq 0 \) (to be verified in
equilibrium) we can represent the diffusion matrix:

$$\sigma_t = \eta \left[ \frac{V_t}{1 - V_t \lambda_t^\top \epsilon_t \lambda_t^\top} + I \right] \sigma_z$$  \hspace{1cm} (51)$$

We then have the following result.

Proposition 1 The REE diffusion matrix \( \sigma_t \) and the variance-covariance matrix \( \Sigma_t \) are non-singular, and

$$\sigma_t^{-1} = \frac{1}{\eta} \sigma_z^{-1} \left[ I - V_t \epsilon_t \lambda_t^\top \right]$$  \hspace{1cm} (52)$$

Proof. By the maintained assumption that \( \sigma_z \) is invertible, the lemma follows directly if we were able to show that \( \left[ \frac{V_t}{\epsilon_t} \epsilon_t \lambda_t^\top + I \right] \) is invertible. From the Sherman-Morrison theorem, this is true if \( 1 + \frac{V_t}{\epsilon_t} \lambda_t^\top \epsilon_t \neq 0 \), which simplifies to \( 1 \neq 0 \). The expression for the inverse is the Sherman-Morrison formula.

4.2 Closed Form Solution

To make further progress in the many asset case, we examine a special case that allows us to solve for the equilibrium in closed form. The special case allows us to reduce the dimensionality of the problem and utilize the ODE solution from the single risky asset case. Our focus here is on the correlation structure of the endogenous returns on the risky assets.

Assumption (Symmetry, S) The diffusion matrix for \( z \) is \( \eta \bar{\sigma}_z I_N \) where \( \bar{\sigma}_z > 0 \) is a scalar and where \( I_N \) is the \( N \times N \) identity matrix. Also, \( \delta^i = \delta \) for all \( i \).

The symmetry assumption enables us to solve the model in closed form and examine the changes in correlation. Together with the i.i.d. feature of the demand shocks we conjecture an REE where \( \mu_t^i = \mu_t^1 \), \( \beta_t^i = \beta_t^1 \), \( \lambda_t^i = \lambda_t^1 \), \( \epsilon_t^i = \epsilon_t^1 \), \( \sigma_t^{ii} = \sigma_t^{11} \) and \( \sigma_t^{ij} = \sigma_t^{12}, i \neq j \). First, notice that \( \epsilon_t \lambda_t^\top = \epsilon_t^1 \lambda_t^1 \mathbf{1}^\top \), and that \( \epsilon_t^\top \lambda_t = N \epsilon_t^1 \lambda_t^1 \), where \( \mathbf{1} \) is a \( N \times 1 \) vector of ones (so that \( \mathbf{1}^\top \) is the \( N \times N \) matrix with the number 1 everywhere).
From (51) we see that the diffusion matrix is given by

$$\eta \sigma_z \left( \frac{V_t \lambda_t^1 \epsilon_t^1}{1 - NV_t \lambda_t^1 \epsilon_t^1} \mathbf{11}^\top + I \right)$$

From here the benefit of symmetry becomes clear. At an REE we only need to solve for one diffusion variable, $\sigma_i^i = \sigma_i^{11}$, since for $i \neq j$ the cross effects $\sigma_i^{ij} = \sigma_i^{12} = \sigma_i^{11} - \eta \tilde{\sigma}_z$ are then determined as well. Recall that $\sigma_i^{ij}$ is the measure of the effect of a change in the demand shock of the $j$th security on the price of the $i$th security, and not the covariance. In other words, it governs the comovements between securities that would otherwise be independent. Define $x_t \equiv x(V_t)$ the solution to the ODE (34) with $\eta$ replaced by $\frac{\eta}{N}$; i.e. $x_t$ is equal to the right-hand-side of (35) with $\eta$ replaced by $\frac{\eta}{N}$. The proof of the following proposition is in the appendix.

**Proposition 2** Assume (S). The following is an REE.

The REE diffusion coefficients are $\sigma_i^i = x_t + \frac{N-1}{N} \eta \tilde{\sigma}_z$, and for $i \neq j$, $\sigma_i^{ij} = x_t - \frac{1}{N} \eta \tilde{\sigma}_z$. Also, $\Sigma_{ii} = \text{Var}_t(\text{return on security } i) = \eta^2 \tilde{\sigma}_z^2 + \frac{1}{N} \left(N^2 x_t^2 - \eta^2 \tilde{\sigma}_z^2\right)$, and for $i \neq j$, $\Sigma_{ij} = \text{Cov}_t(\text{return on security } i, \text{ return on security } j) = \frac{1}{N} \left(N^2 x_t^2 - \eta^2 \tilde{\sigma}_z^2\right)$ and $\text{Corr}_t(\text{return on security } i, \text{ return on security } j) = \frac{1}{N} \left(N^2 x_t^2 - \eta^2 \tilde{\sigma}_z^2\right)$.

Risky holdings are $D_t^i = \frac{V_t}{\alpha N x_t^2}$.

The risk-reward relationship is given by

$$\frac{\mu_t^i - r}{x_t} = \frac{1}{2 \alpha \frac{\alpha}{N} \hat{\sigma}_z} \left\{ 2 \alpha (r^* - r) + \alpha \left(N x_t^2 + \eta^2 \tilde{\sigma}_z^2 - N - \frac{1}{N}\right) - \eta \sqrt{N} \tilde{\sigma}_z + \sqrt{N} \left(x_t - \frac{\eta}{N} \tilde{\sigma}_z\right) \left[2 \alpha^2 r + \frac{\alpha^2 \delta}{V_t} - 2\right] \right\}$$

The intuition and form of the drift term is very similar to the $N = 1$ case and reduces to it if $N$ is set equal to 1.

With multiple securities and with active banks, each idiosyncratic shock is transmitted through the system through the banks’ portfolio decisions. On the one hand this means that less than the full impact of the shock
on security $i$ will be transmitted into the asset return $i$, potentially leading to a less volatile return. The reason is that a smaller fraction of the asset portfolio is invested in asset $i$, reducing the extent of the feedback effect. On the other hand, the demand shocks to assets other than $i$ will be impounded into return $i$, potentially leading to a more volatile return, depending on the extent of mutual cancellations due to the diversification effect on the FIs’ equity. In a world with multiple risky securities satisfying the assumptions in the proposition, the extent of contagion across securities is given by $\sigma_{i}^{ij} = x_t - \frac{1}{\sqrt{N}}\eta\tilde{\sigma}_z$, for $i \neq j$. In the absence of FIs, $x_t = x(0) = \frac{1}{\sqrt{N}}\eta\tilde{\sigma}_z$, so any given security return is unaffected by the idiosyncratic shocks hitting other securities.

For comparison purposes, denote the scalar diffusion coefficient from the $N = 1$ case, as given by (35), by $\sigma_t^{N=1}$. The first direct effect can be characterized as follows: $\sigma_t^{N=1} < \sigma_t^{N=1} \iff \eta\tilde{\sigma}_z < \sigma_t^{N=1}$. In words, each security return is affected less by its own noise term than in a setting with only this one security, for small levels of capital. The reason for this latter effect lies in the fact that any given amount of FI capital needs to be allocated across multiple securities now. For capital levels larger than the critical level $V^* : \sigma_t^{N=1}(V^*) = \eta\tilde{\sigma}_z$, the direct effect is larger than in the $N = 1$ economy because the (now less constrained) risk-neutral FIs tend to absorb aggregate return risk as opposed to idiosyncratic return risk. Whereas all uncertainty vanishes in the $N = 1$ case since FIs insure the residual demand when capital becomes plentiful ($\lim_{V \to \infty} \sigma_t^{N=1} = 0$), with $N > 1$ on the other hand individual volatility remains ($\lim_{V \to \infty} \Sigma_t^{11} = \frac{N-1}{N} \eta^2 \tilde{\sigma}_z^2 > 0$) but the fact that correlations tend to $-1$ means that $\lim_{V \to \infty} \text{Var}(\text{return on the equilibrium portfolio}) = 0$. So again as FI capital increases, aggregate equilibrium return uncertainty is washed out, even though returns continue to have idiosyncratic noise.

Combining direct and indirect effects, return variance is lower in the multi

\footnote{For instance, as $V \to \infty$, we have $\lim_{V \to \infty} \sigma_t^{11} = \frac{N-1}{N} \eta^2 \tilde{\sigma}_z^2 > 0 = \lim_{V \to \infty} \sigma_t^{N=1}$.}
security case if \( V \) is small: \( \Sigma_{i1}^{11} < (\sigma_i^{N=1})^2 \) iff \( \eta^2 \tilde{\sigma}_z^2 / N^2 < x_i^2 \). Still, as in the \( N = 1 \) case securities returns are more volatile with active banks \( (V_t > 0) \), provided capital is not too large.

Diversification across the \( N \) i.i.d. demand shocks lessens the feedback effect on prices to some extent. Since the VaR constraints bind hard for small levels of capital, the fact that idiosyncratic shocks are mixed and affect all securities implies that asset returns become more correlated for small capital levels. FIs tend to raise covariances by allowing the i.i.d. shocks that affect security \( i \) to be also affecting security \( j \neq i \) through their portfolio choices. This effect has some similarities to the wealth effect on portfolio choice described by Kyle and Xiong (2001). The intuition is as follows. Without FIs, returns on all securities are independent. With a binding VaR constraint, in the face of losses, FIs’ risk appetite decreases and they are forced to scale down the risk they have on their books. This leads to joint downward pressure on all risky securities.

This effect is indeed confirmed in an REE, leading to positively correlated returns. This effect is consistent with anecdotal evidence on the loss of diversification benefits suffered by hedge funds and other traders who rely on correlation patterns, when traders are hit by market shocks. The argument
also works in reverse: as FIs start from a tiny capital basis that does not allow them to be much of a player and accumulate more capital, they are eager to purchase high Sharpe ratio securities. This joint buying tends to raise prices in tandem.

Figure 8 shows the correlation as a function of $V$. As can be seen on Figure 8, variances move together, and so do variances with correlations. This echoes the findings in Andersen et al (2001) who show that

“there is a systematic tendency for the variances to move together, and for the correlations among the different stocks to be high/low when the variances for the underlying stocks are high/low, and when the correlations among the other stocks are also high/low.”

They conjecture that these co-movements occur in a manner broadly consistent with a latent factor structure (the $x$ process in our model).
5 Further Results

The logic of the feedback effects that underlies the shapes of the volatility, risk premia and Sharpe Ratio graphs naturally has a number of powerful corollaries that tie in with empirical regularities in the financial markets.

5.1 Leverage Effect

The “leverage effect” refers to the empirical regularity noted by Black (1976) and Schwert (1989) that declining asset prices lead to increased future volatility. Recent work by Kim and Kon (1994), Tauchen, Zhang and Liu (1996) and Anderson, Bollerslev, Diebold and Ebens (2001) find that the leverage effect is stronger for indices than for individual securities. This has been considered as a puzzle for a literal interpretation of the leverage effect, and we are not aware of theoretical explanations for this asymmetry. Our model also finds that the overall market volatility reacts more to falls than individual securities, and our story provides a natural intuition for the effect. As equity $V$ is reduced, prices fall and volatilities increase. Since correlations also increase as equity falls, the volatility of the market portfolio increases more than the volatility of the individual securities underlying the market portfolio. Define $\tilde{V}$ so that $\frac{\partial x_t}{\partial V} (\tilde{V}) = 0$, meaning that the region where equity satisfies $V > \tilde{V}$ corresponds to the usual region right of the hump where capital losses lead to more volatile returns.

Proposition 3 Assume $N > 1$. A decrease in $V$ raises the volatility of the market more than it raises the volatility of an individual constituent security iff $V_t > \tilde{V}$.

Proof. In our model it can easily be verified that the variance of the market portfolio is equal to $\Sigma^m_t := N x_t^2$, and that $\Sigma^{ii}_t = N x_t^2 + \frac{N-1}{N} \eta^2 \sigma_z^2$. A few manipulations verify

$$\frac{\partial \sqrt{\Sigma^m}_t}{\partial V} - \frac{\partial \sqrt{\Sigma^{ii}}_t}{\partial V} = N x_t \frac{\partial x_t}{\partial V} \left[ \frac{1}{\sqrt{\Sigma^m}_t} - \frac{1}{\sqrt{\Sigma^{ii}}_t} \right] < 0$$
iff $V_t > \tilde{V}$, since $\Sigma_i^{-n} < \Sigma_i^{\mu}$ if $N > 1$.

5.2 Derivatives Pricing Implications

It is well known that the Black-Scholes-Merton implied volatilities exhibit a negative skew in moneyness $K/S$ that is fading with longer time to maturity (see for instance Aït-Sahalia and Lo (1998) for a formal econometric analysis). The usual intuition for the relative over-pricing of out-of-the-money (OTM) puts compared to OTM calls within the Black-Scholes-Merton model relies on the fact that OTM puts offer valuable protection against downside “pain points,” and that such a downside either is expected to occur more frequently than similar upside movements or at least occurs in more volatile environments than would a similar upside movement, or that investors are willing to pay more to protect the downside compared to Black-Scholes. These effects imply a fatter left tail of the risk-neutral returns distribution compared to the Gaussian Black-Scholes model, as can be verified on Figure 9.

Our model provides a simple micro-founded alternative channel through which the observed volatility skew is generated. In our framework the REE volatility function $|\sigma_t|$ largely depends negatively on bank capital $V_t$ (except
for very small values of $V_t$). Capital being random, volatility is stochastic. Since the value of the underlying risky asset (viewed as the overall market index) depends positively on bank capital over a large range of capital levels, but its volatility depends mostly negatively on bank capital, one can expect that the option generated implied volatility skew appears in equilibrium.

If one focuses on the at-the-money (ATM, moneyness of one) across various capital levels, one sees that the ATM implied vols (which would in this model be equivalent to the VIX index or similar) are counter-cyclical. An economy with higher capital levels has a lower VIX, and worsening economic circumstances lead to a higher VIX. This is a well-established empirical fact, so much so that the VIX is also referred to as the “investor fear gauge.”

Figure 10 gives the implied volatility surface arising from our model in $(K/S, \text{maturity})$ space. We see the skew for each maturity, as well as a flattening over longer maturities. The flattening is due to the fact that over a longer horizon bank equity will more likely than not have drifted upwards and further out of the danger zone.
6 Concluding Remarks

We have examined a dynamic rational expectations asset pricing model with a banking sector. Our model has stochastic volatility and with “as if” preferences with the feature that banks act as if their preferences are changing in response to market outcomes. The channel through which risk premiums, volatility and risk capacity are connected are the risk constraints that banks operate under. As risk constraints bind harder, effective risk aversion of the banks also increases. We have argued that this simple story of risk aversion feedbacks captures important features of the cyclical nature of banking and its impact on economy-wide risk premiums.

Our discussion has focused mainly on the positive questions, rather than normative, welfare questions on the appropriate role of financial regulation and other institutional features. We recognize that such normative questions will be even more important going forward, especially in the light of the experiences gained in the financial crisis of 2007-9.
Appendix

Lemma (Properties of the Diffusion Term)

\[ S1 \lim_{\tau \to 0} \sigma(V_i) = \eta \sigma_z \]
\[ S2 \lim_{V_i \to \infty} \sigma(V_i) = 0 \] and \( \lim_{V_i \to -\infty} V_i \sigma(V_i) = \infty \). [Call \( f(V) := \frac{\sigma - \eta \sigma_z}{V_i} \) and notice that \( \lim_{V \to 0} f(V) = \lim_{V \to -\infty} \frac{\partial \sigma}{\partial V} \). Since we know the expression for \( \frac{\partial \sigma}{\partial V} \) by (34), we see that the problem can be transformed into \( \lim f = \lim_{V \to 0} 1 \frac{\sigma - \eta \sigma_z}{V} = 0 \).]

\[ S3 \lim_{\tau \to 0} \frac{\partial \sigma}{\partial \tau} = \frac{\eta \sigma_z}{\alpha^2} \] and \( \lim_{\tau \to -\infty} \frac{\partial^2 \sigma}{\partial \tau^2} = \frac{4\eta \sigma_z}{(\alpha^2 \delta_{ij})^2}. \]

\[ S4 \{ V^* \in \mathbb{R} : \sigma(V^*) = 0 \} \] is a singleton. At \( V^* \), \( \sigma \) is strictly decreasing. [The second observation comes from (34) while the first one comes from the fact that the mapping \( V \to \int_{\frac{2\delta}{\alpha}}^\infty \frac{e^{-u}}{u} du \) is a bijection between \( \mathbb{R}_+ \) and \( \mathbb{R} \), so for each chosen \( c \), there is a unique \( V(c) \) setting \( c - \alpha^2 \delta \eta \sigma_z \int_{\frac{2\delta}{\alpha}}^\infty \frac{e^{-u}}{u} du = 0 \).]

\[ S5 \sigma(V) \) has exactly one minimum and one maximum. The minimum is at \( V' \) s.t. \( \sigma(V') < 0 \). The maximum is at \( V'' \) s.t. \( \sigma(V'') > 0 \).

Proof of Proposition 2 First, we can read off (53) the variables of interest as \( \sigma_{ii} = \eta \hat{\sigma}_z \left( \frac{V_{ii}}{1 - N V_{ii} \epsilon_{ii} \lambda_t} + 1 \right) \) and \( \sigma_{ij} = \sigma_{ji} = \sigma_{11} = \eta \hat{\sigma}_z \frac{V_{i1} \lambda_t}{1 - N V_{i1} \epsilon_{i1} \lambda_t} = \sigma_{i1} - \eta \hat{\sigma}_z \).

Next, we compute the variance-covariance matrix, the square of the diffusion matrix (53):

\[ \Sigma_t = \sigma_t \sigma_t = \eta^2 \hat{\sigma}_z^2 [I_N + m_t I] = \eta^2 \hat{\sigma}_z^2 I_N + g_t I \]
where

\[
m_t := N \left( \frac{V_t \epsilon_t^1 \lambda_t^1}{1 - NV_t \epsilon_t^1 \lambda_t^1} \right)^2 + 2 \frac{V_t \epsilon_t^1 \lambda_t^1}{1 - NV_t \epsilon_t^1 \lambda_t^2}
\]

\[
= \frac{1}{\eta^2 \tilde{\sigma}_z^2} \left( \sigma_{11}^t - \eta \tilde{\sigma}_z \right) \left[ 2 \eta \tilde{\sigma}_z + N \left( \sigma_{11}^t - \eta \tilde{\sigma}_z \right) \right]
\]

\[
g_t := m_t \eta^{-2} \tilde{\sigma}_z^{-2}
\]

where we used the fact that \( \eta \tilde{\sigma}_z \frac{V_t \epsilon_t^1 \lambda_t^1}{1 - NV_t \epsilon_t^1 \lambda_t^1} = \sigma_{11}^t - \eta \tilde{\sigma}_z \). Then insert \( \Sigma_t \) into the reward-to-risk equation \( \Sigma_t \lambda_t^1 1 = \frac{\mu_t^1 - r}{\sqrt{\xi}} 1 \) to get

\[
\sqrt{\xi} \lambda_t^1 \left( \eta \tilde{\sigma}_z + N \left( \sigma_{11}^t - \eta \tilde{\sigma}_z \right) \right)^2 = \mu_t^1 - r.
\]

Next compute \( \xi_t \). By definition, \( \xi_t := (\mu_t^1 - r)^2 1^T \Sigma_t^{-1} 1 \). Since \( 1 + g_t (\eta \tilde{\sigma}_z)^{-2} N \neq 0 \), by the Sherman-Morrison theorem we see that

\[
\Sigma_t^{-1} = \left( \eta \tilde{\sigma}_z \right)^{-2} I - \frac{g_t}{(\eta \tilde{\sigma}_z)^{4} + N (\eta \tilde{\sigma}_z)^2 g_t} 11^T
\]

and therefore that

\[
\xi_t = (\mu_t^1 - r)^2 N (\eta \tilde{\sigma}_z)^{-2} \left[ 1 - \frac{N g_t}{(\eta \tilde{\sigma}_z)^{2} + N g_t} \right]
\]

Inserting the expression for \( \xi_t \) into the expression for \( \lambda_t^1 \) we get, using the fact that \( [\eta \tilde{\sigma}_z + N (\sigma_{11}^t - \eta \tilde{\sigma}_z)]^2 = N g_t + (\eta \tilde{\sigma}_z)^2 \),

\[
\lambda_t^1 = \frac{\iota_{AB}}{\sqrt{N} \left[ \eta \tilde{\sigma}_z + N (\sigma_{11}^t - \eta \tilde{\sigma}_z) \right]}
\]

where \( \iota \) is the sign function, \( A := \mu_t^1 - r \) and \( B := N \sigma_{11}^t - (N - 1) \eta \tilde{\sigma}_z \). Using again the fact that \( [\eta \tilde{\sigma}_z + N (\sigma_{11}^t - \eta \tilde{\sigma}_z)]^2 = N g_t + (\eta \tilde{\sigma}_z)^2 \), we see that

\[
\beta_t^1 = \iota_{AB} \frac{1}{\sqrt{N}} \left[ \eta \tilde{\sigma}_z + N (\sigma_{11}^t - \eta \tilde{\sigma}_z) \right]
\]

By definition of \( \epsilon_t^1 \):

\[
\epsilon_t^1 = \frac{1}{\alpha_2^2 \iota_{AB}} \left[ \frac{1}{\sqrt{N}} \left[ \eta \tilde{\sigma}_z + N (\sigma_{11}^t - \eta \tilde{\sigma}_z) \right] + V_t \sqrt{N} \frac{\partial \sigma_{11}^t}{\partial V_t} \right]
\]

(55)

36
Inserting all these expressions into the equation for \( \sigma_{11}^t \), \( \sigma_{11}^t = \eta \tilde{\sigma}_z \frac{1-(N-1)V_t \lambda t}{1-NV_t \lambda z} \) and defining \( x_t := \frac{1}{N}\eta \tilde{\sigma}_z + (\sigma_{11}^t - \eta \tilde{\sigma}_z) \), the resulting equation is the ODE (34) with \( \eta \) replaced by \( \eta/N \) and where \( \sigma(V) \) is replaced by \( x(V) \).

As to risky holdings, we know that \( \Delta_t = \frac{1}{N} \tilde{\sigma}_z \). Noticing that \( \tilde{\sigma}_z + N(\sigma_{11}^t - \eta \tilde{\sigma}_z) = Nx_t \), we find that \( \lambda_t = \frac{1}{N^2} \) from which the expression for \( D_t \) follows.

Finally we compute the risk premia. Using Itô’s Lemma on (21) we get

\[
\mu_t^1 - \frac{1}{2} \Sigma_t^1 = \alpha \epsilon_t^1 (\text{drift of } V_t) + r^* \, dt + \frac{1}{2} \alpha \frac{\partial \epsilon_t^1}{\partial V_t} (\text{diffusion of } V_t)^2
\]

Now the drift of capital can be seen to be equal to

\[
\text{drift of } V_t = rV_t + D_t^T (\mu_t - r) = rV_t + (\mu_t^1 - r) \frac{V_t}{\alpha \sqrt{Nx_t}}
\]

and using \( \sigma_t = (x_t - \frac{1}{N} \eta \tilde{\sigma}_z) I + \eta \tilde{\sigma}_z I \) the squared diffusion term can be verified to be equal to \( \frac{1}{N} \). The drift equation becomes

\[
(\mu_t^1 - r) \left[ 1 - \epsilon_t^1 \frac{V_t}{\sqrt{Nx_t}} \right] = \frac{1}{2} (\sigma_t^1)^2 + \alpha^2 \epsilon_t^1 rV_t + \frac{1}{2} \alpha \frac{\partial \epsilon_t^1}{\partial V_t}
\]

We can rewrite (55) by inserting the ODE for \( x_t \) to get rid of the partial derivative term:

\[
\epsilon_t^1 = t AB \frac{\sqrt{N}}{V_t} \left( x_t - \frac{\eta}{N} \tilde{\sigma}_z \right)
\]

Performing the differentiation of \( \epsilon_t^1 \) and inserting into the drift equation completes the proof.
References


Adrian, Tobias, Emanuel Moench and Hyun Song Shin (2009) “Financial Intermediation, Asset Prices and Macroeconomic Dynamics” FRB New York Staff Report 422


Rytchkov, Oleg (2008) “Dynamic Margin Constraints,” working paper, University of Texas


