Balance Sheet Capacity and Endogenous Risk*

by Jon Danielsson Hyun Song Shin Jean–Pierre Zigrand

THE PAUL WOOLLEY CENTRE WORKING PAPER SERIES NO 16 DISCUSSION PAPER NO 665

DISCUSSION PAPER SERIES

January 2011

Jon Danielsson is a Reader in Finance at the London School of Economics. His research interests include financial risk modelling, regulation of financial markets, models of extreme market movements, market liquidity, and financial crisis. He has a PhD in the economics of financial markets from Duke University, has published extensively in both academic and practitioner journals, and has presented his work in a number of universities, public institutions, and private firms. Hyun Song Shin is the Hughes-Rogers Professor of Economics at Princeton University. Before coming to Princeton in 2006, he was Professor of Finance at the London School of Economics. His current research interests are in financial economics with particular reference to financial institutions, disclosures, risk and financial stability issues, topics on which he has published widely both in academic and practitioner outlets. He has served as editor or editorial board member of several scholarly journals, and has served in an advisory capacity to central banks and policy organizations on financial stability issues. He is a fellow of the Econometric Society and of the British Academy. Jean-Pierre Zigrand is a Reader in Finance at the London School of Economics. His research interests include asset pricing, financial intermediation theory and general equilibrium theory. Any opinions expressed here are those of the author and not necessarily those of the FMG. The research findings reported in this paper are the result of the independent research of the author and do not necessarily reflect the views of the LSE.

Balance Sheet Capacity and Endogenous Risk^{*}

Jon Danielsson London School of Economics Hyun Song Shin Princeton University

Jean–Pierre Zigrand London School of Economics

> This version: January 2011

Abstract

Banks operating under Value-at-Risk constraints give rise to a welldefined 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 recent financial crisis has served as a reminder of the important role played by banks and other financial intermediaries in driving the financial cycle. The depletion of bank capital and subsequent deleveraging by banks has been a central theme in the discussion of the credit crunch and its impact on the real economy. In this paper, we explore the idea that the banking sector has a well-defined "balance sheet capacity" that encapsulates its ability to take on risky exposures, and that the fluctuations in this capacity is the engine that drives the financial cycle.

The notion of "balance sheet capacity" sits uncomfortably with textbook discussions of how corporate balance sheets are determined. In a world where the Modigliani and Miller (MM) theorems hold, we can separate the decision on the size of the balance sheet (selection of the projects to take on) from the financing of the projects (composition of liabilities in terms of debt and equity). The size of the balance sheet is determined by the set of positive net present value (NPV) projects, which is normally treated as being exogenous. The focus is on the liabilities side of the balance sheet, in determining the relative mix of equity and debt in financing the assets. Even when the conditions for the MM theorems do not hold, the textbook discussion starts with the assets of the firm as given, in order to focus on the financing decision alone.

However, the distinguishing feature of banking sector balance sheets is that they fluctuate widely over the financial cycle. Credit increases rapidly during the boom but increases less rapidly (or even decreases) during the downturn, driven partly by shifts in the banks' willingness to take on risky positions over the cycle. The evidence that banks' willingness to take on risky exposures fluctuates over the cycle is especially clear for financial intermediaries that operate in the capital market. Figure 1, taken from Adrian and Shin (2010), shows the scatter chart of the quarterly change in assets against the quarterly change in leverage of the (then) five stand-alone US investment banks.¹ Leverage is the ratio of total assets to equity.

The first feature to note is that leverage is procyclical in the sense that leverage is high when balance sheets are large, while leverage is low when balance sheets are small. This is the opposite relationship to that for households, whose leverage is high when balance sheets are small. For instance, if a household owns a house that is financed by a mortgage, leverage falls when the house price increases, since the equity of the household is increasing at a much faster rate than assets.

¹Bear Stearns, Goldman Sachs, Lehman Brothers, Merrill Lynch and Morgan Stanley



Figure 1: Leverage Growth and Total Asset Growth (Source: Adrian and Shin (2010))

The vertical axis in Figure 1 measures quarterly asset growth given by the change in log assets. The horizontal axis gives the change in leverage, as measured by the change in log assets minus the change in log equity. Formally, we have:

Asset growth = log
$$D(t + 1) - \log D(t)$$

Leverage Growth = log $D(t + 1) - \log D(t)$
 $- (\log V(t + 1) - \log V(t))$

where D denotes total assets and V denotes equity. The 45-degree line therefore represents the set of points where

$$\log V\left(t+1\right) - \log V\left(t\right) = 0$$

and so represents the constant equity line. Any straight line with slope 1 in Figure 1 represents the set of points where equity is growing at a constant rate, where the growth rate is given by the intercept. In Figure 1, we see that the slope of the scatter chart is close to 1, suggesting that equity is increasing at a constant rate on average. Thus, unlike the textbook discussion of the Modigliani-Miller theorem, it is equity that seems to play the role of the exogenous variable and total assets (the size of the balance sheet) is the endogenous variable that is determined by the willingness of banks to take on risky exposure.

Adrian and Shin (2010) pointed to the risk management policies of financial intermediaries as a possible explanation of the fluctuations in assets while equity is the forcing variable. Suppose that banks aim to keep enough equity capital to meet its overall Value-at-Risk (VaR). If we denote by VaR the value at risk per dollar of assets, then equity capital V must satisfy

$$V = \text{VaR} \times D$$

implying that leverage L satisfies

$$L = D/V = 1/VaR$$

If Value-at-Risk is low in expansions and high in contractions, leverage is high in expansions and low in contractions – that is, leverage is procyclical. This suggests that there is a notion of *balance sheet capacity* for the banking sector that depends on first, the size of its capital base (its equity) and second, the amount of lending that can be supported by each unit of capital. Total assets are then determined by the multiplication of the two.

However, realized risks should itself be endogenous, and depend on the ability of banks to take on risky exposures. When the banking sector suffers depletion of capital due to losses on its assets, its capacity to take on risky exposures diminishes as the Value-at-Risk constraint tightens. In other words, balance sheet capacity, Value-at-Risk and market risk premiums should all be determined simultaneously in equilibrium.

In what follows, we solve for the equilibrium in closed form in a dynamic banking model and examine how balance sheet capacity, volatility and risk premiums are jointly determined.

One key feature of the equilibrium is that risk premiums are high when banking sector capital is depleted, implying that projects that previously received funding from the banking sector no longer do so with depleted capital. This is a result that is reminiscent of a "credit crunch" due to banking sector losses, and follows from the following confluence of forces. Banks are risk neutral but their capacity to take on risky exposures is limited by their capital cushion. As their capital is depleted, their Value-at-Risk constraints bind harder, and their behavior resembles that of risk-averse investors. Indeed, the Lagrange multiplier associated with the capital constraint enters into the banks' lending decisions just like a risk aversion parameter. As banks suffer erosion of their capital, equilibrium volatility increases at the same time as their "as if" risk aversion also increases. 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 no longer receives funding.

As well as capturing the features of a credit crunch, the predictions from our model are also consistent with accmulating empirical evidence that the balance sheet capacity of banks and other financial intermediaries predict returns on various risky assets (we discuss the literature shortly). Faster asset growth today is associated with lower future risky asset returns, consistent with lower risk premiums. The common thread is the fact that intermediary balance sheet capacity holds useful information on risk premiums, and hence on the future expected returns on risky assets.

The closed-form solution offered in this paper offers a number of useful insights. For the case with a single risky asset, the closed-form solution for the equilibrium can be written in terms of the relative size of the banking sector in the financial system as a whole. Conveniently, our model also has the feature that the distribution of bank capital does not matter, and only the aggregate banking sector capital matters in equilibrium. Thus, there is a well-defined notion of an economy-wide banking sector lending capacity.

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.

For the case of many risky assets, 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 technical issues such as the volatility of volatility as well as topics in derivatives pricing, such as the shape of the volatility curve.

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 Gennotte 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. Examples 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) and Rytchkov (2008). Brunnermeier and Pedersen (2009) emphasize the "margin spirals" that result where capital constraints set off amplified feedback Garleanu and Pedersen (2009) extend the CAPM by incorporateffects. ing a capital constraint to show how assets with the same fundamental risk may trade at different prices. He and Krishnamurthy (2007) have studied 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) shares with our paper the focus on financial intermediation although their focus is on examining the effect on the real economy through real investment decisions.

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 fluctuations 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 commodities 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 associated 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. To our knowledge, our paper is the first to solve the fixed point problem in closed form, yielding tractable expressions for equilibrium returns, volatility and correlations. Moreover, 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 2 depicts the relationships. Under this interpretation, the market value of the risky securities can be thought of as the marked-to-market



Figure 2: Intermediated and Directly Granted Credit

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_t^i . We will look for an equilibrium in which the price processes for the risky assets follow:

$$\frac{dP_t^i}{P_t^i} = \mu_t^i dt + \sigma_t^i dW_t \qquad ; i = 1, \dots, N$$
(1)

where W_t is an $N \times 1$ vector of independent Brownian motions, and where the scalar μ_t^i and the $1 \times N$ vector σ_t^i 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,² but it 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 θ_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

$$D_t^i := \theta_t^i P_t^i \tag{2}$$

The budget constraint of the trader is

$$b_t B_t = V_t - \theta_t^{\top} P_t = V_t - \sum_{i=1}^N D_t^i$$
 (3)

where V_t is the trader's capital and where x^{\top} is the transpose of x. The dynamic budget constraint governs the evolution of capital in the usual way:

$$dV_t = \theta_t^\top dP_t + b_t dB_t$$

= $\left[rV_t + D_t^\top (\mu_t - r) \right] dt + D_t^\top \sigma_t dW_t$ (4)

where D^{\top} denotes the transpose of D, and where σ_t is the $N \times N$ diffusion matrix, row i of which is σ_t^i . In (4), 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 (4), the expected capital gain is

$$E_t[dV_t] = [rV_t + D_t^{\top}(\mu_t - r)]dt$$
(5)

and the variance of the trader's equity is

$$\operatorname{Var}_t(dV_t) = D_t^{\top} \sigma_t \sigma_t^{\top} D_t dt \tag{6}$$

²See 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.

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

$$\Sigma_t := \sigma_t \sigma_t^\top \tag{7}$$

The bank is risk-neutral, and maximizes return (5) subject to its VaR constraint, where VaR is some constant α 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 α 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:

$$\max_{D_t} rV_t + D_t^{\top}(\mu_t - r) \qquad \text{subject to} \qquad \alpha \sqrt{D_t^{\top} \sigma_t \sigma_t^{\top} D_t} \le V_t \qquad (8)$$

Once the dollar values $\{D_t^i\}_{i=1}^N$ of the risky assets are determined, the bank's residual bond holding is determined by the balance sheet identity:

$$b_t B_t = V_t - \sum_i D_t^i \tag{9}$$

The first-order condition for the optimal D is

$$\mu_t - r = \alpha (D_t^\top \Sigma_t D_t)^{-1/2} \gamma_t \Sigma_t D_t$$
(10)

where γ_t is the Lagrange multiplier associated with the VaR constraint. Hence,

$$D_{t} = \frac{1}{\alpha (D_{t}^{\top} \Sigma_{t} D_{t})^{-1/2} \gamma_{t}} \Sigma_{t}^{-1} (\mu_{t} - r)$$
(11)

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,

$$V_t = \alpha \sqrt{D_t^{\top} \Sigma_t D_t} \tag{12}$$

and therefore

$$D_t = \frac{V_t}{\alpha^2 \gamma_t} \Sigma_t^{-1} (\mu_t - r)$$
(13)

Notice that the optimal portfolio is similar to the mean-variance optimal portfolio allocation, where the Lagrange multiplier γ_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 riskneutral, but they will behave like risk averse investors whose risk aversion appears to shift in line with the Lagrange multiplier γ . Substituting into (12) and rearranging we have

$$\gamma_t = \frac{\sqrt{\xi_t}}{\alpha} \tag{14}$$

where

$$\xi_t := (\mu_t - r)^{\top} \Sigma_t^{-1} (\mu_t - r) \ge 0$$
(15)

The Lagrange multiplier γ_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 α becomes small, the VaR constraint binds less and banks' willingness to take on risk increases.

Notice that the Lagrange multiplier γ_t does not depend directly on equity V_t . Intuitively, an additional unit of capital relaxes the VaR constraint by a multiple α of standard deviation, leading to an increase in the expected return equal to a multiple α 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 (15).

Finally, we can solve for the risky asset holdings as

$$D_t = \frac{V_t}{\alpha \sqrt{\xi_t}} \Sigma_t^{-1} (\mu_t - r) \tag{16}$$

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 (16) 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 downwardsloping 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, $y_t = (y_t^1, \dots, y_t^N)$ where

$$y_t = \Sigma_t^{-1} \begin{bmatrix} \delta^1 \left(z_t^1 - \ln P_t^1 \right) \\ \vdots \\ \delta^N \left(z_t^N - \ln P_t^N \right) \end{bmatrix}$$
(17)

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 δ^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 δ 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

$$\frac{V_t}{\alpha\sqrt{\xi_t}}(\mu_t - r) + \begin{bmatrix} \delta^1 \left(z_t^1 - \ln P_t^1\right) \\ \vdots \\ \delta^N \left(z_t^N - \ln P_t^N\right) \end{bmatrix} = 0$$
(18)

Equilibrium prices are therefore

$$P_t^i = \exp\left(\frac{V_t}{\alpha\delta^i\sqrt{\xi_t}}(\mu_t^i - r) + z_t^i\right); \qquad i = 1,\dots,N$$
(19)

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:

$$\frac{dP_t^i}{P_t^i} = \mu_t^i dt + \sigma_t^i dW_t \qquad ; i = 1, \dots, N$$
(20)

where W_t is an $N \times 1$ vector of independent Brownian motions, and where the scalar μ_t^i and $1 \times N$ vector σ_t^i 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:

$$dz_t^i = r^* dt + \eta \sigma_z^i dW_t \tag{21}$$

where σ_z^i is a 1 × N vector that governs which Brownian shocks will get impounded into the demand shocks and therefore govern the correlation structure of the demand shocks. We assume that the stacked $N \times N$ matrix σ_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 $\{\sigma_t^i\}$ of the return process depends on the (exogenous) shock terms $\{\sigma_z^i\}$, 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:

$$\frac{dP_t}{P_t} = \mu_t dt + \sigma_t dW_t \tag{22}$$

where μ_t and σ_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:

$$dz_t = r^* dt + \eta \sigma_z dW_t \tag{23}$$

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

$$\xi_t = \frac{(\mu_t - r)^2}{\sigma_t^2} \tag{24}$$

Substituting into (19), 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

$$P_t = \exp\left(z_t + \frac{\sigma_t V_t}{\alpha \delta}\right) \tag{25}$$

From (22) we have, by hypothesis,

$$d\ln P_t = \left(\mu_t - \frac{1}{2}\sigma_t^2\right)dt + \sigma_t dW_t \tag{26}$$

Meanwhile, taking the log of (25) and applying Itô's Lemma gives

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

= $r^* dt + \eta \sigma_z dW_t + \frac{1}{\alpha \delta} d(\sigma_t V_t)$
= $r^* dt + \eta \sigma_z dW_t + \frac{1}{\alpha \delta} (\sigma_t dV_t + V_t d\sigma_t + dV_t d\sigma_t)$ (27)

Now use Itô's Lemma on $\sigma(V_t)$:

$$d\sigma_t = \frac{\partial\sigma}{\partial V_t} dV_t + \frac{1}{2} \frac{\partial^2\sigma}{\partial (V_t)^2} (dV_t)^2$$
$$= \left\{ \frac{\partial\sigma}{\partial V_t} \left[rV_t + \frac{V_t(\mu_t - r)}{\alpha\sigma_t} \right] + \frac{1}{2} \frac{\partial^2\sigma}{\partial (V_t)^2} \left(\frac{V_t}{\alpha}\right)^2 \right\} dt + \frac{\partial\sigma}{\partial V_t} \frac{V_t}{\alpha} dW_t \quad (28)$$

where (28) follows from

$$dV_t = [rV_t + D_t(\mu_t - r)]dt + D_t\sigma_t dW_t$$
$$= \left[rV_t + \frac{V_t(\mu_t - r)}{\alpha\sigma_t}\right]dt + \frac{V_t}{\alpha}dW_t$$
(29)

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 σ_t itself.

Substituting back into (27) and regrouping all dt terms into a new drift term:

$$d\ln P_t = (\text{drift term}) \ dt + \left[\eta \sigma_z + \frac{1}{\alpha \delta} \left(\sigma_t \frac{V_t}{\alpha} + V_t \frac{\partial \sigma_t}{\partial V_t} \frac{V_t}{\alpha}\right)\right] dW_t$$
(30)

We can solve for the equilibrium diffusion σ_t by comparing coefficients between (30) and (26). We have an equation for the equilibrium diffusion given by:

$$\sigma(V_t) = \eta \sigma_z + \frac{1}{\alpha \delta} \left(\sigma_t \frac{V_t}{\alpha} + V_t \frac{\partial \sigma_t}{\partial V_t} \frac{V_t}{\alpha} \right)$$
(31)

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

$$V_t^2 \frac{\partial \sigma}{\partial V_t} = \alpha^2 \delta(\sigma_t - \eta \sigma_z) - V_t \sigma_t \tag{32}$$

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

$$\sigma(V_t) = \frac{1}{V_t} e^{-\frac{\alpha^2 \delta}{V_t}} \left[c - \alpha^2 \delta \eta \sigma_z \int_{-\frac{\alpha^2 \delta}{V_t}}^{\infty} \frac{e^{-u}}{u} du \right]$$
(33)

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

$$\sigma(V_t) = \eta \sigma_z \frac{\alpha^2 \delta}{V_t} \exp\left\{-\frac{\alpha^2 \delta}{V_t}\right\} \times \operatorname{Ei}\left(\frac{\alpha^2 \delta}{V_t}\right)$$
(34)

where Ei(w) is the well-known exponential integral function:

$$\operatorname{Ei}(w) \equiv -\int_{-w}^{\infty} \frac{e^{-u}}{u} du \tag{35}$$

The Ei (w) function is defined provided $w \neq 0$. The expression $\alpha^2 \delta/V_t$ which appears prominently in the closed form solution (34) can be interpreted as the relative scale or size of the value investor sector (parameter δ) compared to the banking sector (total capital V_t normalized by VaR).

The closed form solution also reveals much about the basic shape of the volatility function $\sigma(V_t)$. Consider the limiting case when the banking sector is very small, that is, $V_t \to 0$. Then $\alpha^2 \delta/V_t$ becomes large, but the exponential term exp $\{-\alpha^2 \delta/V_t\}$ dominates, and the product of the two goes to zero. 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 $\eta \sigma_z$. As we will illustrate with the numerical example to be presented below, the role of the Ei (w) term is to tie down the end point so that the limiting volatility is given by this fundamental volatility.

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

$$\mu_t = r + \frac{\sigma_t}{2\alpha\eta\sigma_z} \left\{ 2\alpha(r^* - r) + \alpha\sigma_t^2 - \eta\sigma_z + (\sigma_t - \eta\sigma_z) \left[2\alpha^2 r + \frac{\alpha^2\delta}{V_t} - 2 \right] \right\}$$
(36)

We can see that μ_t depends on the diffusion σ_t , so that when the expression in the square brackets is positive, μ_t is increasing in σ_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 σ_t , the tighter is the risk constraint, and hence the higher is the expected return μ_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 $\mu_t - \frac{(\eta\sigma_z)^2}{2} = r^*$, confirming our interpretation of r^* as the value investor sector's benchmark log-return.

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

$$\frac{\mu_t - r}{\sigma_t} = \frac{1}{2\alpha\eta\sigma_z} \left\{ 2\alpha(r^* - r) + \alpha\sigma_t^2 - \eta\sigma_z + (\sigma_t - \eta\sigma_z) \left[2\alpha^2 r + \frac{\alpha^2\delta}{V_t} - 2 \right] \right\}$$
(37)

The countercyclical shape of the Sharpe Ratio follows directly from the shape of the diffusion coefficient σ_t .

3.1 Numerical Example

We illustrate the properties of our closed form solution by means of a numerical example. Figure 3 plots the equilibrium diffusion σ_t and the drift μ_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 (34), the plot of σ_t is nonmonotonic, 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 σ_z and η (= 0.3 × 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.³

What Figure 3 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 σ_t and μ_t are small. At the opposite extreme, if V_t

 $^{^{3}}$ See Mele (2007) for a discussion of the stylized facts, and for a model generating countercyclical statistics in a more standard framework.



Figure 3: Risk Premium and Volatility as Functions of Bank Equity



Figure 4: Lagrange Multiplier of Bank Capital Constraint

is very large, then banks begin to act more and more like an unconstrained trader. Since the trader is risk-neutral, the expected drift μ_t is pushed down to the risk-free rate, and the volatility σ_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 σ_t . Due to the risk constraint, the riskneutral banks behave "as if" they were risk averse, and the equilibrium drift μ_t reflects this feature of the model. The risk premium μ_t rises with σ_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 4. We see that the Sharpe ratio rises and falls roughly the same pattern with σ_t and μ_t . However, the notable feature of Figure 4 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 4 shows is that the increased leverage may start to come into play for large values of V.

Figure 5 gives scatter charts for the relationship between the asset growth and leverage for four sample realizations of the model. These relationships are the theoretical counterparts generated in our model to the empirical relationship depicted in Figure 1 for the Wall Street investment banks. 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. 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 5 shows this best.



Figure 5: Scatter Charts of Asset and Leverage Changes

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:

$$\frac{dP_t^i}{P_t^i} = \mu_t^i dt + \sigma_t^i dW_t \tag{38}$$

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

$$dz_t^i = r^* dt + \eta \sigma_z^i dW_t \tag{39}$$

where σ_z^i 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 μ and σ respectively. For notational convenience, we define the scaled reward-to-risk factor

$$\lambda_t := \frac{1}{\sqrt{\xi_t}} \Sigma_t^{-1} (\mu_t - r) \tag{40}$$

Also, we use the following shorthands:

$$\beta_t^i := \frac{1}{\sqrt{\xi_t}} (\mu_t^i - r) \tag{41}$$

$$\epsilon_t^i := \frac{1}{\alpha^2 \delta^i} \beta_t^i + \frac{V_t}{\alpha^2 \delta^i} \frac{\partial \beta_t^i}{\partial V_t}$$
(42)

and where

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

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 (19) we have:

$$\sigma_t^i = \tilde{\epsilon}_t^i V_t \tilde{\lambda}_t^\top \tilde{\sigma}_t + \eta \sigma_z^i \tag{43}$$

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

We can solve the fixed point problem by specifying a beliefs updating process $(\tilde{\mu}, \tilde{\sigma})$ 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:

$$\begin{bmatrix} \tilde{\mu}_t \\ \tilde{\sigma}_t \end{bmatrix} = \begin{bmatrix} \mu_t(\tilde{\mu}_t, \tilde{\sigma}_t) \\ \sigma_t(\tilde{\mu}_t, \tilde{\sigma}_t) \end{bmatrix}.$$
(44)

By stacking into a diffusion matrix, at a REE the diffusion matrix satisfies

$$\sigma_t = V_t \epsilon_t \lambda_t^\top \sigma_t + \eta \sigma_z \tag{45}$$

Using the fact that $\lambda_t^{\top} \sigma_t \sigma_t^{\top} = \beta_t^{\top}$, σ_t satisfies the following matrix quadratic equation $\sigma_t \sigma_t^{\top} = \eta \sigma_z \sigma_t^{\top} + V_t \epsilon_t \beta_t^{\top}$ so that

$$(\sigma_t - \eta \sigma_z) \sigma_t^\top = V_t \epsilon_t \beta_t^\top \tag{46}$$

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 (43) (keeping in mind that $\frac{V_t}{\alpha}\lambda_t\sigma_t$ equals the diffusion term of equity)

$$\sigma_t^i = \underbrace{\left(\frac{1}{\alpha\delta^i}\beta_t^i\right)}_{\text{feedback effect on vol}} (\text{vol of capital}) + \underbrace{\left(\frac{V}{\alpha\delta^i}\frac{\partial\beta_t^i}{\partial V_t}\right)}_{\text{feedback effect on vol}} (\text{vol of capital}) + \eta\sigma_z^i$$

We now solve for a representation of σ_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

$$e_t := 1 - V_t \lambda_t^{\top} \epsilon_t$$

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

$$\sigma_t = \eta \left[\frac{V_t}{1 - V_t \lambda_t^\top \epsilon_t} \epsilon_t \lambda_t^\top + I \right] \sigma_z \tag{47}$$

We then have the following result.

Proposition 1 The REE diffusion matrix σ_t and the variance-covariance matrix Σ_t are non-singular, and

$$\sigma_t^{-1} = \frac{1}{\eta} \sigma_z^{-1} \left[I - V_t \epsilon_t \lambda_t^\top \right]$$
(48)

Proof. By the maintained assumption that σ_z is invertible, the lemma follows directly if we were able to show that $\begin{bmatrix} V_t \\ e_t \\ e_t \\ k_t \\ \lambda_t^\top \\ e_t \end{bmatrix}$ is invertible. From the Sherman-Morrison theorem, this is true if $1 + \frac{V_t}{e_t} \lambda_t^\top \\ \epsilon_t \\ \epsilon_t$

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 $\eta \tilde{\sigma}_z I_N$ where $\tilde{\sigma}_z > 0$ 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 $\epsilon_t \lambda_t^\top = \epsilon_t^1 \lambda_t^1 \mathbf{11}^\top$, and that $\epsilon_t^\top \lambda_t = N \epsilon_t^1 \lambda_t^1$, where **1** is a $N \times 1$ vector of ones (so that $\mathbf{11}^\top$ is the $N \times N$ matrix with the number 1 everywhere).

From (47) we see that the diffusion matrix is given by

$$\eta \sigma_z \left(\frac{V_t \lambda_t^1 \epsilon_t^1}{1 - N V_t \lambda_t^1 \epsilon_t^1} \mathbf{1} \mathbf{1}^\top + I \right)$$
(49)

From here the benefit of symmetry becomes clear. At an REE we only need to solve for one diffusion variable, $\sigma_t^{ii} = \sigma_t^{11}$, since for $i \neq j$ the cross effects $\sigma_t^{ij} = \sigma_t^{12} = \sigma_t^{11} - \eta \tilde{\sigma}_z$ are then determined as well. Recall that σ_t^{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 by $x_t \equiv x(V_t)$ the solution to the ODE (32) with η replaced by $\frac{\eta}{N}$, i.e. x_t is equal to the right-hand-side of (33) with η 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 $\sigma_t^{ii} = x_t + \frac{N-1}{N}\eta\tilde{\sigma}_z$, and for $i \neq j$, $\sigma_t^{ij} = x_t - \frac{1}{N}\eta\tilde{\sigma}_z$. Also, $\Sigma_t^{ii} = Var_t(return on security i) = \eta^2\tilde{\sigma}_z^2 + \frac{1}{N}(N^2x_t^2 - \eta^2\tilde{\sigma}_z^2)$, and for $i \neq j$, $\Sigma_t^{ij} = Cov_t(return on security i, return on security j) = \frac{1}{N}(N^2x_t^2 - \eta^2\tilde{\sigma}_z^2)$ and $Corr_t(return on security i, return on security j) = \frac{Nx_t^2 - \frac{1}{N}\eta^2\tilde{\sigma}_z^2}{Nx_t^2 + \frac{N-1}{N}\eta^2\tilde{\sigma}_z^2}$.

Risky holdings are $D_t^i = \frac{V_t}{\alpha N^{3/2} x_t}$. The risk-reward relationship is given by

$$\frac{\mu_t^i - r}{x_t} = \frac{1}{2\alpha \frac{\eta}{N}\tilde{\sigma}_z} \left\{ 2\alpha(r^* - r) + \alpha \left(Nx_t^2 + \eta^2 \tilde{\sigma}_z^2 \frac{N - 1}{N} \right) - \frac{\eta}{\sqrt{N}} \tilde{\sigma}_z + \sqrt{N} \left(x_t - \frac{\eta}{N} \tilde{\sigma}_z \right) \left[2\alpha^2 r + \frac{\alpha^2 \delta}{V_t} - 2 \right] \right\}$$
(50)

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 $\sigma_t^{ij} = x_t - \frac{1}{N}\eta\tilde{\sigma}_z$, for $i \neq j$. In the absence of FIs, $x_t = x(0) = \frac{1}{N}\eta\tilde{\sigma}_z$, so any given security return is unaffected by the idiosyncratic shocks hitting other securities.



Figure 6: Cross Contagion across Risky Assets

For comparison purposes, denote the scalar diffusion coefficient from the N = 1 case, as given by (33), by $\sigma_t^{N=1}$. The first direct effect can be characterized as follows: $\sigma_t^{11} < \sigma_t^{N=1}$ 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 economy⁴ 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 $\lim_{V \to \infty} Var(return on the equilibrium portfolio)= 0$. 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 security case if V is small: $\Sigma_t^{11} < (\sigma_t^{N=1})^2$ iff $\eta^2 \tilde{\sigma}_z^2 / N^2 < x_t^2$. 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

⁴For instance, as $V \to \infty$, we have $\lim_{V \to \infty} \sigma_t^{11} = \frac{N-1}{N} \eta \tilde{\sigma}_z > 0 = \lim_{V \to \infty} \sigma_t^{N=1}$.



Figure 7: Return Correlations across Risky Assets

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 7 shows the correlation as a function of V. As can be seen on Figure 7, 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 $\frac{\partial x_t}{\partial V_t}(\tilde{V}) = 0$, 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 $\Sigma_t^m := N x_t^2$, and that $\Sigma_t^{ii} = N x_t^2 + \frac{N-1}{N} \eta^2 \tilde{\sigma}_z^2$. A few manipulations verify

$$\frac{\partial \sqrt{\Sigma_t^m}}{\partial V} - \frac{\partial \sqrt{\Sigma_t^{ii}}}{\partial V} = N x_t \frac{\partial x_t}{\partial V} \left[\frac{1}{\sqrt{\Sigma_t^m}} - \frac{1}{\sqrt{\Sigma_t^{ii}}} \right] < 0$$

iff $V_t > \tilde{V}$, since $\Sigma_t^m < \Sigma_t^{ii}$ if N > 1.



Figure 8: Skewed Density over Outcomes

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 8.

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



Figure 9: Implied Volatility Surface

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 9 gives the implied volatility surface arising from our model in (K/S, 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)

 $[\mathbf{S1}] \lim_{V_t \to 0} \sigma(V_t) = \eta \sigma_z$

[S2] $\lim_{V_t\to\infty} \sigma(V_t) = 0$ and $\lim_{V_t\to\infty} V_t \sigma(V_t) = \infty$ **[S3]** $\lim_{V_t\to0} \frac{\partial\sigma}{\partial V_t} = \frac{\eta\sigma_z}{\alpha^2\delta}$ and $\lim_{V_t\to0} \frac{\partial^2\sigma}{\partial V_t^2} = \frac{4\eta\sigma_z}{(\alpha^2\delta)^2}$. [Call $f(V) := \frac{\sigma_t - \eta\sigma_z}{V_t}$ and notice that $\lim_{V\to0} f(V) = \lim_{V\to0} \frac{\partial\sigma}{\partial V}$. Since we know the expression for $\frac{\partial\sigma}{\partial V}$ by (32), we see that the problem can be transformed into $\lim f = \lim \frac{1}{V_t} [\alpha^2 \delta f(V) - \sigma]$. In turn, we can replace $\frac{\sigma_t}{V_t}$ definitionally by $f + \frac{\eta\sigma_z}{V}$ to get to $\lim f = \lim \frac{f(V)[\alpha^2\delta - V] - \eta\sigma_z}{V}$. If $\lim f$ is not equal to the constant given here, then the RHS diverges. Since the denominator of the RHS converges to zero, so must the numerator. Thus the constant is the one shown here. The proof of the second limit is similar.]

[S4] { $V^* \in \mathbb{R} : \sigma(V^*) = 0$ } is a singleton. At V^* , σ is strictly decreasing. [The second observation comes from (32) while the first one comes from the fact that the mapping $V \mapsto \int_{-\frac{\alpha^2 \delta}{V}}^{\infty} \frac{e^{-u}}{u} du$ is a bijection between \mathbb{R}_+ and \mathbb{R} , so for each chosen c, there is a unique V(c) setting $\left[c - \alpha^2 \delta \eta \sigma_z \int_{-\frac{\alpha^2 \delta}{V(c)}}^{\infty} \frac{e^{-u}}{u} du\right] = 0$.]

[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 (49) the variables of interest as $\sigma_t^{ii} = \eta \tilde{\sigma}_z \left(\frac{V_t \epsilon_t \lambda_t}{1 - N V_t \epsilon_t \lambda_t} + 1 \right)$ and $\sigma_t^{ij} = \sigma_t^{12} = \eta \tilde{\sigma}_z \frac{V_t \epsilon_t \lambda_t}{1 - N V_t \epsilon_t \lambda_t} = \sigma_t^{11} - \eta \tilde{\sigma}_z$.

Next, we compute the variance-covariance matrix, the square of the diffusion matrix (49):

$$\Sigma_t = \sigma_t \sigma_t = \eta^2 \tilde{\sigma}_z^2 [I_N + m_t \mathbf{1} \mathbf{1}^\top] = \eta^2 \tilde{\sigma}_z^2 I_N + g_t \mathbf{1} \mathbf{1}^\top$$

where

$$m_t := N \left(\frac{V_t \epsilon_t^1 \lambda_t^1}{1 - N V_t \epsilon_t^1 \lambda_t^1} \right)^2 + 2 \frac{V_t \epsilon_t^1 \lambda_t^1}{1 - N V_t \epsilon_t^1 \lambda_t^2}$$
$$= \frac{1}{\eta^2 \tilde{\sigma}_z^2} (\sigma_t^{11} - \eta \tilde{\sigma}_z) \left[2\eta \tilde{\sigma}_z + N (\sigma_t^{11} - \eta \tilde{\sigma}_z) \right]$$
$$g_t := m_t \eta^2 \tilde{\sigma}_z^2$$

where we used the fact that $\eta \tilde{\sigma}_z \frac{V_t \epsilon_t^1 \lambda_t^1}{1 - NV_t \epsilon_t^1 \lambda_t^1} = \sigma_t^{11} - \eta \tilde{\sigma}_z$. Then insert Σ_t into the reward-to-risk equation $\Sigma_t \lambda_t^1 \mathbf{1} = \frac{\mu_t^1 - r}{\sqrt{\xi}} \mathbf{1}$ to get $\sqrt{\xi_t} \lambda_t^1 \left[\eta \tilde{\sigma}_z + N(\sigma_t^{11} - \eta \tilde{\sigma}_z)\right]^2 = \mu_t^1 - r$.

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

$$\Sigma_t^{-1} = (\eta \tilde{\sigma}_z)^{-2} I - \frac{g_t}{(\eta \tilde{\sigma}_z)^4 + N(\eta \tilde{\sigma}_z)^2 g_t} \mathbf{1} \mathbf{1}^\top$$

and therefore that

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

Inserting the expression for ξ_t into the expression for λ_t^1 we get, using the fact that $[\eta \tilde{\sigma}_z + N(\sigma_t^{11} - \eta \tilde{\sigma}_z)]^2 = Ng_t + (\eta \tilde{\sigma}_z)^2$,

$$\lambda_t^1 = \frac{\iota_A \iota_B}{\sqrt{N} [\eta \tilde{\sigma}_z + N(\sigma_t^{11} - \eta \tilde{\sigma}_z)]}$$

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

$$\beta_t^1 = \iota_A \iota_B \frac{1}{\sqrt{N}} \left[\eta \tilde{\sigma}_z + N (\sigma_t^{11} - \eta \tilde{\sigma}_z) \right]$$

By definition of ϵ_t^1 :

$$\epsilon_t^1 = \frac{1}{\alpha^2 \delta} \iota_A \iota_B \left[\frac{1}{\sqrt{N}} [\eta \tilde{\sigma}_z + N(\sigma_t^{11} - \eta \tilde{\sigma}_z)] + V_t \sqrt{N} \frac{\partial \sigma_t^{11}}{\partial V_t} \right]$$
(51)

Inserting all these expressions into the equation for σ_t^{11} , $\sigma_t^{11} = \eta \tilde{\sigma}_z \frac{1-(N-1)V_t \lambda_t^1 \epsilon_t^1}{1-NV_t \lambda_t^1 \epsilon_t^1}$ and defining $x_t := \frac{1}{N} \eta \tilde{\sigma}_z + (\sigma_t^{11} - \eta \tilde{\sigma}_z)$, the resulting equation is the ODE (32) with η replaced by η/N and where $\sigma(V)$ is replaced by x(V).

As to risky holdings, we know that $D_t = \frac{V_t}{\alpha} \lambda_t$. Noticing that $\eta \tilde{\sigma}_z + N(\sigma_t^{ii} - \eta \tilde{\sigma}_z) = N x_t$, we find that $\lambda_t^1 = \frac{1}{N^{3/2} x_t}$ from which the expression for D_t follows.

Finally we compute the risk premia. Using Itô's Lemma on (19) we get

$$\mu_t^i - \frac{1}{2}\Sigma_t^{11} = \alpha \epsilon_t^1 (\text{drift of } V_t) + r^* dt + \frac{1}{2} \alpha \frac{\partial \epsilon_t^1}{\partial V_t} (\text{diffusion of } V_t)^2$$

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

drift of
$$V_t = rV_t + D_t^{\top}(\mu_t - r) = rV_t + (\mu_t^1 - r)\frac{V_t}{\alpha\sqrt{N}x_t}$$

and using $\sigma_t = \left(x_t - \frac{1}{N}\eta\tilde{\sigma}_z\right)\mathbf{1}\mathbf{1}^\top + \eta\tilde{\sigma}_z I$ the squared diffusion term can be verified to be equal to $\frac{V_t^2}{\alpha^2}$. The drift equation becomes

$$(\mu_t^1 - r) \left[1 - \epsilon_t^1 \frac{V_t}{\sqrt{N}x_t} \right] = \frac{1}{2} (\sigma_t^{11})^2 + \alpha^2 \epsilon_t^1 r V_t + \frac{1}{2} \frac{V_t^2}{\alpha} \frac{\partial \epsilon_t^1}{\partial V_t}$$

We can rewrite (51) by inserting the ODE for x_t to get rid of the partial derivative term:

$$\epsilon_t^1 = \iota_A \iota_B \frac{\sqrt{N}}{V_t} \left(x_t - \frac{\eta}{N} \tilde{\sigma}_z \right)$$

Performing the differentiation of ϵ^1_t and inserting into the drift equation completes the proof.

References

Adrian, Tobias, Erkko Etula and Hyun Song Shin (2009) "Risk Appetite and Exchange Rates" FRB New York Staff Report 361

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

Adrian, Tobias and Hyun Song Shin (2008) "Financial Intermediary Leverage and Value at Risk," working paper, Federal Reserve Bank of New York and Princeton University. Federal Reserve Bank of New York Staff Reports, 338.

Adrian, Tobias and Hyun Song Shin (2010) "Liquidity and Leverage," *Journal of Financial Intermediation* 19 (3), 418-437

Aït-Sahalia, Yacine and Andrew Lo (1998), "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," Journal of Finance, 53, 499-547.

Aiyagari, S. Rao and Mark Gertler (1999) ""Overreaction" of Asset Prices in General Equilibrium," Review of Economic Dynamics 2, pp. 3-35.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Ebens, H., (2001) "The Distribution of Stock Return Volatility," Journal of Financial Economics, 61, 1, 43-76.

Basak, Suleyman, and Benjamin Croitoru (2000) "Equilibrium mispricing in a capital market with portfolio constraints," The Review of Financial Studies 13, pp 715-748.

Black, Fischer (1976) "Studies of stock price volatility changes," Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section, pp 177-181.

Black, Fischer and Myron Scholes (1973) "The Pricing of Options and Corporate Liabilities," Journal of Political Economy, 81, 637-654.

Bernardo, Antonio E. and Ivo Welch, (2004) "Financial Market Runs," Quarterly Journal of Economics, 119, 135-158.

Brunnermeier, Markus (2008) "De-Ciphering the Credit Crisis of 2007," forthcoming, Journal of Economic Perspectives.

Brunnermeier, Markus and Lasse Pedersen (2009) "Market Liquidity and Funding Liquidity," Review of Financial Studies, 22, 2201-2238.

Brunnermeier, Markus and Yuliy Sannikov (2009) "A Macroeconomic Model with a Financial Sector" working paper, Princeton University

Chabakauri, Georgy (2008) "Asset Pricing in General Equilibrium with Constraints," working paper, London Business School

Corsi, Fulvio and Uta Kretschmer, Stefan Mittnik, and Christian Pigorsch (2005), "The Volatility of Realized Volatility," Working Paper No. 2005/33, Center for Financial Studies, Frankfurt

Coudert, Virginie and Mathieu Gexa (2008), "Does risk aversion drive financial crises? Testing the predictive power of empirical indicators," Journal of Empirical Finance Volume 15, Issue 2, March 2008, Pages 167-184

Danielsson, Jon, Paul Embrechts, Charles Goodhart, Felix Muennich, Con Keating, Olivier Renault and Hyun Song Shin (2001) "An Academic Response to Basel II" Financial Markets Group Special Paper 130 http://hyunsongshin.org/www/basel2.pdf

Danielsson, Jon, Hyun Song Shin (2003) "Endogenous Risk" in *Modern Risk Management: A History*, Risk Books, 2003

Danielsson, Jon, Hyun Song Shin and Jean-Pierre Zigrand (2004) "The Impact of Risk Regulation on Price Dynamics" *Journal of Banking and Finance*, 28, 1069-1087

Danielsson, Jon and Jean-Pierre Zigrand (2008) "Equilibrium Asset Pricing with Systemic Risk", *Economic Theory*, 35, 293-319

Deutsche Bundesbank (2005), "Risk appetite in a dynamic financial market environment," Monthly Report, October 2005.

Etula, Erkko (2009) "Risk Appetite and Commodity Returns" working paper, FRB New York.

Gai, Prasanna and Nicholas Vause (2005), "Measuring investors' risk appetite", International Journal of Central Banking 2, pp. 167-118.

Garleanu, Nicolae and Lasse Heje Pedersen (2009) "Margin-Based Asset Pricing and Deviations from the Law of One Price," working paper. Geanakoplos, John (1997) "Promises, Promises" In W.B. Arthur, S. Durlauf and D. Lane (eds.), The Economy as an Evolving Complex System, II. Reading MA: Addison-Wesley, 1997, pp. 285-320.

Geanakoplos, John (2009) "The Leverage Cycle" NBER Macroeconomics Annual 2009.

Geanakoplos, John and William Zame (2003) "Liquidity, Default, and Crashes: Endogenous Contracts in General Equilibrium," Advances in Economics and Econometrics: Theory and Applications, Eighth World Conference, Volume II, Econometric Society Monographs (2003), pp. 170-205.

Gennotte, Gerard and Hayne Leland (1990) "Hedging and Crashes", American Economic Review, 999–1021.

Gromb, Denis and Dimitri Vayanos (2002) "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs," Journal of Financial Economics 66, pp. 361-407.

He, Zhiguo and Arvind Krishnamurthy (2007) "Intermediary Asset Pricing," Northwestern University.

Heston, S. (1993), "A closed-form solution for options with stochastic volatility with applications to bond and currency options," Review of Financial Studies 6, 327-343.

Jones, Christopher (2003), "The dynamics of stochastic volatility: evidence from underlying and options markets," Journal of Econometrics, 116, 181-224.

Kashyap, Anil, Raghuram Rajan and Jeremy Stein (2008) "Rethinking Capital Regulation" paper prepared for the Federal Reserve Bank of Kansas City Symposium at Jackson Hole, 2008.

Kim, D. and Kon, S.J. (1994), "Alternative Models for the Conditional Heteroskedasticity of Stock Returns," Journal of Business, 67, 563-598

Kiyotaki, N and John Moore (1997) "Credit Cycles" Journal of Political Economy, 105, 211 - 248

Kyle, A. S. and Xiong, Wei (2001) "Contagion as a Wealth Effect" Journal of Finance, 56, 1401 - 1440.

Holmström, Bengt, and Jean Tirole (1998) "Private and Public Supply of Liquidity," Journal of Political Economy 106, pp. 1-40.

Mele, Antonio (2007), "Asymmetric Stock Market Volatility and the Cyclical Behavior of Expected Returns," Journal of Financial Economics 86, 446-478 (2007)

Merton, Robert (1973), "Theory of Rational Option Pricing," Bell Journal of Economics and Management Science, 4, 141-183.

Morris, Stephen and Hyun Song Shin (2004) "Liquidity Black Holes" Review of Finance, 8, 1-18

Oehmke, Martin (2008) "Liquidating Illiquid Collateral" working paper, Columbia University GSB

Pavlova, Anna and Roberto Rigobon (2008) "The Role of Portfolio Constraints in the International Propagation of Shocks," Review of Economic Studies 75, 1215-1256

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

Schwert, G (1989) "Why Does Stock Market Volatility Change Over Time?," Journal of Finance 44, 1115, 1153.

Sherman, J. and W. J. Morrison (1949) "Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix," Ann. Math. Statist., 20, 621.

Shleifer, Andrei and Robert Vishny (1997) "The Limits of Arbitrage" Journal of Finance, 52, 35-55.

Tauchen, G., Zhang, H. and Liu, M., 1996, "Volume, Volatility and Leverage: A Dynamic Analysis," Journal of Econometrics, 74, 177-208

Xiong, Wei (2001) "Convergence Trading with Wealth Effects: An Amplification Mechanism in Financial Markets", Journal of Financial Economics 62, 247-292