Risk Appetite and Endogenous Risk*

Jon Danielsson                   Hyun Song Shin
London School of Economics       Princeton University

Jean–Pierre Zigrand
London School of Economics

February 14, 2009

Abstract

Market volatility reflects traders’ actions, while their actions depend on perceptions of risk. Equilibrium volatility is the fixed point of the mapping that takes perceived risk to actual risk. We solve for equilibrium stochastic volatility in a dynamic setting where risk-neutral traders operate under Value-at-Risk constraints. We derive a closed form solution for the stochastic volatility function in the benchmark model with a single risky asset. Even though the underlying fundamental risks remain constant, the resulting dynamics generate stochastic volatility through traders’ reactions in equilibrium. Volatilities, expected returns and Sharpe ratios are shown to be countercyclical. If the purpose of financial regulation is to shield the financial system from collapse, then basing regulation on individually optimal risk management may not be enough.

*We are grateful to Markus Brunnermeier, Antonio Mele, Anna Pavlova, Dimitri Vayanos, Wei Xiong and participants at the Adam Smith asset pricing conference and at MIT for comments on an earlier draft.
1 Introduction

Financial crises are often accompanied by large price changes, but large price changes by themselves do not constitute a crisis. Public announcements of important macroeconomic statistics, such as the U.S. employment report, are sometimes marked by large, discrete price changes at the time of announcement. However, such price changes are arguably the signs of a smoothly functioning market that is able to incorporate new information quickly. The market typically finds composure quite rapidly after such discrete price changes.

In contrast, the distinguishing feature of crisis episodes is that they seem to gather momentum from the endogenous responses of the market participants themselves. Rather like a tropical storm over a warm sea, they appear to gather more energy as they develop. As financial conditions worsen, the willingness of market participants to bear risk seemingly evaporates. They curtail their exposures and generally attempt to take on a more prudent, conservative stance. However, the shedding of exposures results in negative spillovers on other market participants from the sale of assets or withdrawal of credit. As prices fall or measured risks rise or previous correlations break down (or some combination of the three), market participants respond by further cutting exposures. The global financial crisis of 2007–8 has served as a live laboratory for many such distress episodes.

The main theme of our paper is the endogeneity of risk. The risks impacting financial markets are attributable (at least in part) to the actions of market participants. In turn, market participants’ actions depend on perceived risk. In equilibrium, risk should be understood as the fixed point of the mapping that maps perceived risk to actual risk. In what follows, we solve a dynamic asset pricing model where equilibrium risk is derived as such a fixed point. Put differently, our task in this paper is to solve a stochastic volatility model, where the stochastic volatility function is solved as a fixed point of the mapping that takes conjectured volatility functions to realized volatility functions. The equilibrium stochastic volatility is the “endogenous risk” referred to in the title of this paper.

One purpose of developing a model of endogenous risk is so that we can study the propagation of financial booms and distress, and to identify and quantify the amplification channels through which such effects operate. Among other things, we can make precise the notion that market participants appear to become “more risk-averse” in response to deteriorating market out-
comes. For economists, preferences and beliefs would normally be considered as being independent of one another.

However, we can distinguish “risk appetite” which motivates traders’ actions, from “risk aversion”, which is a preference parameter hard-wired into agents’ characteristics. A trader’s risk appetite may change even if his preferences are unchanged. The reason is that risk taking may be curtailed by the constraints that traders operate under, such as those based on Value-at-Risk (VaR). In our dynamic asset pricing model, all active traders are risk-neutral, but they operate under Value-at-Risk (VaR) constraints. The Lagrange multiplier associated with the VaR constraint is the key quantity in our model. It plays two important roles. First, it affects the portfolio choice of the traders. Second, we show that the Lagrange multiplier is related to a generalized Sharpe ratio for the set of risky assets traded in the market as a whole, and hence depends on the forecast probability density over future outcomes. Through the Lagrange multiplier, beliefs and risk appetite are thus linked. To an outside observer, it would appear that market participants’ preferences change with minute-by-minute changes in market outcomes. Crucially, shocks may be amplified through the feedback effects that operate from volatile outcomes to reduced capacity to bear risk. In this sense, the distinction between “risk appetite” and “risk aversion” is more than a semantic quibble. This distinction helps us understand how booms and crises play out in the financial system.

In the benchmark case where there is a single risky asset, we are able to solve the equilibrium stochastic volatility function in closed form. We do this by deriving an ordinary differential equation for the diffusion term for the price of the risky asset that must be satisfied by all fixed points of the equilibrium mapping that takes perceived risk to actual risk. Fortunately, this ordinary differential equation can be solved in closed form, giving us a closed form solution for endogeneous risk.

Although our intended contribution is primarily theoretical, our solution also reveals several suggestive features that are consistent with empirical properties of asset returns found in practice. For instance, even when the stochastic shocks that hit the underlying fundamentals of the risky assets are i.i.d., the resulting equilibrium dynamics exhibit time-varying and stochastic volatility. Furthermore, option-implied volatilities, as well as volatilities of volatilities, are shown to be countercyclical in our equilibrium solution, consistent with the empirical evidence. Our model also generates countercyclical and convex (forward looking, rational) risk premia and Sharpe ratios. In a
more general version of our model with multiple risky assets, we show that the increased risk aversion and higher volatilities coincides with increased correlations in equilibrium, even though the underlying shocks are constant, and independent across the risky assets.

By pointing to the endogenous nature of risk, we highlight the role played by risk management rules used by active market participants which serve to amplify aggregate fluctuations. Although it is a truism that ensuring the soundness of each individual institution ensures the soundness of the financial system, this proposition is vulnerable to the fallacy of composition. Actions that an individual institution takes to enhance its soundness may undermine the soundness of others. If the purpose of financial regulation is to shield the financial system from collapse, it is not enough to base financial regulation on the “best practice” of individually optimal risk management policies, as is done under the current Basel II capital regulations.

While our model has suggestive features that are consistent with the empirical evidence, we also recognize the limitations of our model in serving as a framework that can be used directly for empirical work. The reason is that our model has just one state variable (the total equity capital of the active traders), and so cannot accommodate important features such as history-dependence. For instance, we cannot accommodate the notion that a long period of tranquil market conditions serve to accumulate vulnerabilities in the financial system that are suddenly exposed when the financial cycle turns (arguably, an important feature of the global financial crisis of 2007-8). There is a rich seam of future research that awaits further work which develops our model further in order to do justice to such empirical questions.

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) literature review.

1.1 Related Literature

Crisis dynamics and liquidity issues were studied by Genotte and Leland (1990), Genakoplos (1997) and Geanakoplos and Zame (2003) who provided theoretical approaches based on competitive equilibrium analysis, and the
informational role of prices in a rational expectations equilibrium. Shleifer and Vishny’s (1997) observation that margin constraints limit the ability of arbitrageurs to exploit price differences, as well as Holmstrom and Tirole’s (1997) work on debt capacities brought ideas and tools from corporate finance into the study of financial market fluctuations.

Building on these themes has been a spate of recent theoretical work. 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. Xiong (2001) is one of the few papers that examines the fixed point of the equilibrium correspondence, and solves it numerically. 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.

More closely related to our work are papers where endogenous balance sheet constraints enter as an additional channel of contagion to the pure wealth effect. Examples include Aiyagari and Gertler (1999), Basak and Croitoru (2000), Gromb and Vayanos (2002), Brunnermeier (2008) and Brunnermeier and Pedersen (2007), Chabakauri (2008) and Rychkov (2008). 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, and the co-movement of equity prices across countries increases in periods of binding 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.

Our incremental contribution to the wealth effect literature is to incorporate risk-based constraints on active traders, and thereby endogenize risk and risk appetite simultaneously. Relative to the other papers with time-varying risk-bearing capacity, our incremental contribution is to solve for the equilibrium stochastic volatility function as a fixed point of the mapping that takes the conjectured stochastic volatility function to the realized stochastic volatility function. Equilibrium stochastic volatility exhibit many features, such as the countercyclical volatility as reflected in the “smirk” in option-implied volatility, as well as countercyclical correlation and Sharpe ratios.

More directly, this paper builds on our earlier work on Lagrange multiplier associated with Value-at-risk constraints (Danielsson, Shin and Zigrand (2004), Danielsson and Zigrand (2008)). In this earlier work, we dealt with
backward-looking learning rather than solving for equilibrium in a rational
have also emphasized the role of fluctuating Lagrange multipliers associated
with balance sheet constraints in determining risk-bearing capacity. 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 con-
structed 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.

Risk amplification through Value-at-Risk constraints pose questions on
the limits of a regulatory system that relies solely on the “best practice” of
individually prudent risk management rules. The current Basel II capital
requirements rests on such a philosophy, but our paper is one in a long series
of recent papers mentioned above that shows the externalities that exist
between financial institutions, and which point to the importance of taking
a system-wide perspective when considering rules for financial regulations
(see Danielsson et al. (2001), Morris and Shin (2008) and Brunnermeier et
al. (2009)).

Our wish-list of stylized facts of aggregate market fluctuations (such as
countercyclical volatility, risk premia and Sharpe ratios) has been established
for stocks, as summarized for instance in Campbell (2003). For instance,
Black (1976) and Schwert (1989) observed that wealth variability falls as
showed that expected returns fall as wealth rises. While intuitive, it is not
obvious that volatility and risk-premia move together. For instance, Abel
(1988) shows in a general equilibrium model that volatility and risk-premia
vary together only if the coefficient of relative risk aversion is less than one.
Black (1990) provides a setup with one risky security in which only a rela-
tive risk aversion parameter larger than one can lead both to countercyclical
volatility and risk premia, and at the same time lead to a solution of the eq-
uity premium puzzle and the consumption smoothing puzzle (via coefficients
of relative risk aversion of the felicity function and of the value function
that differ). There is a large literature attempting to match better some
of these stylized facts using consumption-based models extended in various
directions, see for instance the habit formation models by Campbell and
Cochrane (1999) and by Chan and Kogan (2002), the behavioural model of

6
Barberis, Huang and Santos (2001) and the limited stock market participation model of Basak and Cuoco (1998). By contrast, our approach is based less on the consumption side of passive consumers (“Main Street”) than on the actions of risk controlled financial institutions (“Wall Street”).

2 The Model

Let time be indexed by $t \in [0, \infty)$. There are $N > 0$ non-dividend paying risky securities as well as a risk-free security. We will focus later on the case where $N = 1$, but we state the problem for the general $N$ risky 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:

$$\frac{dP_i^t}{P_i^t} = \mu_i^t dt + \sigma_i^t dW_t \quad ; i = 1, \ldots, N \tag{1}$$

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 coefficients that will be solved in equilibrium. The risk-free security has price $B_t$ at date $t$, which is given exogenously by $B_0 = 1$ and $dB_t = rB_t dt$, where $r$ is constant.

Our model has two types of traders - the active traders (the financial institutions, or “FIs”) and passive traders. The passive traders play the role of supplying exogenous downward-sloping demand curves for the risky assets, which form the backdrop for the activities of the active traders. As mentioned above, this modeling strategy has been successfully adopted in the recent literature, and we will also pursue the strategy here. In solving for our equilibrium price dynamics, our focus is on the active traders. The price dynamics in equilibrium will be solved as a rational expectations equilibrium with respect to the active traders’ beliefs. We first state the portfolio choice problem of the active traders.

2.1 Portfolio Choice Problem of Active Trader

The active traders (the FIs) are assumed to be short-horizon traders who maximize the instantaneous expected returns on their portfolio. But each trader is subject to a risk constraint where his capital $V$ is sufficiently large to cover his Value-at-Risk (VaR). We use “capital” and “equity” interchangeably in what follows.
It is beyond the scope of our paper to provide microfoundations for the VaR rule\textsuperscript{1}, but it would not be an exaggeration to say that capital budgeting practices based on measured risks (such as VaR) have become pervasive among institutions at the heart of the global financial system. Such practices have also been encouraged by the regulators, through the Basel II bank capital rules. The short-horizon feature 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.

Let $\theta_{it}$ be the number of units of the $i$th risky securities held at date $t$, and denote the dollar amount invested in risky security $i$ by

$$D_{it} := \theta_{it} P_{it} \quad (2)$$

The budget constraint of the trader is

$$b_{it} B_{it} = V_{it} - \theta_{it} \cdot P_{it} \quad (3)$$

where $V_{it}$ is the trader’s capital. The “self-financingness” condition governs the evolution of capital in the usual way.

$$dV_{it} = \theta_{it} \cdot dP_{it} + b_{it} dB_{it}$$

$$= [rV_{it} + D_{i}^{\top} (\mu_{it} - r)] dt + D_{i}^{\top} \sigma_{i} dW_{it} \quad (4)$$

where $D^\top$ denotes the transpose of $D$, and where $\sigma_{i}$ is the $N \times N$ diffusion matrix, row $i$ of which is $\sigma_{i}^t$. 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

$$E_{it}[dV_{it}] = [rV_{it} + D_{i}^{\top} (\mu_{it} - r)] dt \quad (5)$$

and the variance of the trader’s equity is

$$\text{Var}_{it}(dV_{it}) = D_{i}^{\top} \sigma_{i} \sigma_{i}^{\top} D_{i} dt \quad (6)$$

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

$$\Sigma_{t} := \sigma_{i} \sigma_{i}^{\top} \quad (7)$$

\textsuperscript{1}See Adrian and Shin (2008) for one possible microfoundation in a contacting model with moral hazard.
The trader is risk-neutral, and maximizes the capital gain (5) subject to his VaR constraint, where VaR is $\alpha$ times the forward-looking standard deviation of returns on equity. The number $\alpha$ is just a normalizing constant, and does not enter materially into the analysis. The equity $V_t$ is the state variable for the trader. Assuming that the trader is solvent (i.e. $V_t > 0$), the maximization problem can be written as:

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

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

$$b_t B_t = V_t - \sum_i D_t^i \quad (9)$$

The first-order condition for the optimal $D_t$ is

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

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

$$D_t = \frac{1}{\alpha (D_t^\top \Sigma_t D_t)^{-1/2} \gamma_t} \Sigma_t^{-1}(\mu_t - r) \quad (11)$$

When $\mu_t \neq r$, the objective function is monotonic in $D_t$ by risk-neutrality, and the constraint must bind. Hence,

$$V_t = \alpha \sqrt{D_t^\top \Sigma_t D_t} \quad (12)$$

and therefore

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

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 traders in our model are risk-neutral, but they will behave like risk averse traders whose risk aversion appears to shift in line with the Lagrange multiplier $\gamma$. Substituting into (12) and rearranging we have

$$\gamma_t = \frac{\sqrt{\xi_t}}{\alpha} \quad (14)$$
where
\[ \xi_t := (\mu_t - r)^\top \Sigma_t^{-1} (\mu_t - r) \geq 0 \] (15)

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 traders’ risk appetite 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 (15).

Finally, we can solve for the risky asset holdings as
\[ D_t = \frac{V_t}{\alpha \sqrt{\xi_t}} \Sigma_t^{-1} (\mu_t - r) \] (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 Amplification of Shocks

The fact that optimal holding of risky assets is homogeneous of degree one in equity \( V_t \) is a crucial channel for the propagation of shocks through feedback trading. It is easiest to see the impact of the optimal portfolio rule through a concrete example. Consider a trader whose balance sheet is given as follows.

<table>
<thead>
<tr>
<th>Assets ( \sum_i D_t^i )</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_t )</td>
<td>( -b_t B_t )</td>
</tr>
</tbody>
</table>

In this example, the trader is leveraged. He finances the holding \( \sum_i D_t^i \) of risky securities with equity \( V_t \) and debt of \( -b_t B_t \). The two sides of the
balance sheet add up due to (9). The leverage of the trader (the ratio of total assets to equity) can then be obtained from (16) and is given by

\[
\frac{1^T \Sigma_t^{-1}(\mu_t - r)}{\alpha \sqrt{\xi_t}}
\]

where \(1^T\) is the \(1 \times N\) unit row vector consisting of 1’s.

Due to leverage, the trader’s demand response to price shocks amplifies price shocks. To see this, suppose that the trader holds just one risky security and (for illustration) suppose that the trader aims to maintain a constant leverage ratio of 10. The initial balance sheet has 100 dollars’ worth of the security on the asset side, which has been funded with equity of 10 and debt of 90.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Securities, 100</td>
<td>Equity, 10</td>
</tr>
<tr>
<td></td>
<td>Debt, 90</td>
</tr>
</tbody>
</table>

Suppose there is shock to the price of the security which raises the value by 1% to 101. Assuming that the value of debt is unchanged, the burden of adjustment falls on the equity, which has to increase by 10% to 11.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Securities, 101</td>
<td>Equity, 11</td>
</tr>
<tr>
<td></td>
<td>Debt, 90</td>
</tr>
</tbody>
</table>

Leverage then falls to \(101/11 = 9.18\). If the trader targets leverage of 10, he takes on additional debt of \(D\) to purchase \(D\) worth of securities on the asset side so that

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

The solution is \(D = 9\). Thus, an increase in the price of the security leads to an additional purchase of the security. The demand response is therefore upward-sloping. If the greater demand for the risky security puts further upward pressure on its price, then there is the potential for a feedback effect where an initial price shock leads to greater demand for the security,
which raises the price, which increases demand further. There is an analogous potential feedback effect following a negative shock to the price of the security.

In this simple numerical example, we have assumed a constant leverage ratio of 10, but the amplifying effect would be operative as long as the target leverage (17) does not fall too much following a price shock.\(^2\) Also, the extent of the amplification of the initial shock will depend on the price elasticity of the risky security. If increased demand for the risky security puts large upward pressure on the price of the risky security, then the feedback effect will be strong. In the rest of the paper, we demonstrate that the amplification of initial shocks to prices sketched above is a key channel through which risk becomes endogenous.

### 2.3 Closing the Model with Passive Traders

We close the model by introducing passive traders who supply exogenous, downward-sloping demand curves for the risky assets. The slope of the passive traders’ demand curves will determine the size of the price feedback effect sketched above, and will be a parameter in our analysis. We will assume that the passive traders in aggregate have the following vector-valued exogenous demand schedule for the risky assets, \(y_t = (y_{1t}, \ldots, y_{Nt})\) where

\[
y_t = \sum_{i=1}^{N} \delta_i \left( rt + \eta z_{it} - \ln P_i^t \right)
\]

where \(P_i^t\) is the market price for risky security \(i\) and where \(z_{it}\) is a positive demand shock to the demand of asset \(i\) (or a negative supply shock to security \(i\)) to be specified further. Each demand curve can simply be viewed as a downward sloping demand hit by demand shocks, with \(\delta_i\) 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. The market-clearing condition \(D_t + y_t = 0\) can be written as

\(^2\)In another context, Adrian and Shin (2007) show that leverage of the (then) five major US investment banks is actually procyclical - i.e. that leverage increases when balance sheets expand.
\[
\frac{V_t}{\alpha \sqrt{\xi_t}} (\mu_t - r) + \begin{bmatrix}
\delta^1 (rt + \eta z^1_t - \ln P^1_t) \\
\vdots \\
\delta^N (rt + \eta z^N_t - \ln P^N_t)
\end{bmatrix} = 0 
\] (19)

Equilibrium prices are therefore

\[
P^i_t = \exp \left( \frac{V_t}{\alpha \delta^i \sqrt{\xi_t}} (\mu^i_t - r) + rt + \eta z^i_t \right); \quad i = 1, \ldots, N
\] (20)

### 2.4 Solution Strategy

We proceed to solve for the rational expectations equilibrium (REE) of our model. Our strategy in solving for an equilibrium is to begin with some exogenous stochastic process that drives the passive traders’ demands for the risky assets (the “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 follow the processes:

\[
\frac{dP^i_t}{P^i_t} = \mu^i_t dt + \sigma^i_t dW_t \quad ; i = 1, \ldots, N
\] (21)

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 passive traders:

\[
dz^i_t = \sigma^i_z dW_t
\] (22)

where \( \sigma^i_z \) is a \( 1 \times N \) 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 \( \sigma_z \) is of full rank.

Our focus is on the way that the (endogenous) diffusion terms \( \{\sigma^i_t\} \) 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.
We will examine the general problem with $N$ risky assets in a later section, but we first look at the case of a single risky asset. In this case, we can obtain a closed form solution.

### 3 Equilibrium with Single Risky Asset

Consider the case with a single risky asset. We will look for an equilibrium where the price of the risky asset follows the process:

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

(23)

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 passive trader’s demands:

$$dz_t = \sigma_z dW_t$$

(24)

where $\sigma_z > 0$ is a known constant. For the single risky asset case, note that

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

(25)

Substituting into (20), 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

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

(26)

From (23) we have, by hypothesis,

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

(27)

Meanwhile, taking the log of (26) and applying Itô’s Lemma gives

$$d \ln P_t = dt \left( rt + \eta z_t + \frac{\sigma_t V_t}{\alpha \delta} \right) = \eta d z_t + \frac{1}{\alpha \delta} d(\sigma_t V_t)$$

$$= \eta \sigma_z dW_t + \frac{1}{\alpha \delta} (\sigma_t dV_t + V_t d\sigma_t + dV_t d\sigma_t)$$

(28)
Now use Itô’s Lemma on $\sigma(V_t)$:

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

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

where (29) follows from

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

and the fact that $D_t = \frac{V_t}{\alpha \sigma_t}$. Notice that $(dV_t)^2 = \left( \frac{V_t}{\alpha} \right)^2 dt$. Substituting back into (28) and regrouping all $dt$ terms into a new drift term:

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

We can solve for the equilibrium diffusion $\sigma_t$ by comparing coefficients between (30) and (27). We have an equation for the equilibrium diffusion given by:

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

which can be written as the ordinary differential equation:

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

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

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

where $c$ is an arbitrary constant of integration. We thus obtain a closed form solution to the rational expectations equilibrium for the single risky asset case.

Note the multiplicity of equilibria. There is one distinct rational expectations equilibrium for each choice of $c$. This is so even though the structure
of the REE is similar across the choice of the scaling factor $c$. We return to discuss the role of $c$ later in our paper.

The equilibrium drift $\mu_t$ (the expected instantaneous return on the risky asset) can be solved analogously, and is given by

$$\mu_t = r + \frac{\sigma_t}{2\alpha\eta} \left\{ \alpha\sigma_t^2 - \eta \sigma_z + (\sigma_t - \eta \sigma_z) \left[ 2\alpha^2 r + \frac{\alpha^2 \delta}{V_t} - 2 \right] \right\} \tag{34}$$

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 traders 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. The information contained in the equilibrium drift $\mu_t$ and its relationship with the diffusion $\sigma_t$ can be summarized better in the expression for the Sharpe ratio, which is:

$$\frac{\mu_t - r}{\sigma_t} = \frac{1}{2\alpha\eta} \left\{ \alpha\sigma_t^2 - \eta \sigma_z + (\sigma_t - \eta \sigma_z) \left[ 2\alpha^2 r + \frac{\alpha^2 \delta}{V_t} - 2 \right] \right\} \tag{35}$$

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

### 3.1 Properties of Equilibrium

We illustrate the properties of our model graphically. Figure 1 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 $r = 0.01$, $\delta = 1$, $\alpha = 3$, $\sigma_z = 0.2$, $\eta = 1$ and $c = 10$.

Note that $\sigma_t$ is non-monotonic, with a peak when $V_t$ is low.

We can understand this pattern in terms of the amplifying demand response discussed earlier. In figure 3, we plot the portfolio of the trader as a function of $V_t$, and note that as $V_t$ increases, the trader takes an increasingly large position in the risky security. When the holding of the bond $b_tB_t$ turns negative, the

---

3 Also see Mele (2007) for a discussion of the stylized facts, and for a model generating countercyclical statistics in a more standard framework.
interpretation is that the trader is leveraged, and is financing his long position in the risky security with debt. Then, as discussed above, the demand response of the trader to price changes amplifies initial price shocks. When there is a positive shock to the price, equity increases at a faster rate than total assets, so that leverage falls, and the VaR constraint becomes slack. The trader than loads up on the long position. In other words, a positive shock to price leads to greater purchases of the security, putting further upward pressure on the price.

What figure 1 reveals is that such a feedback effect 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 active trader’s purchase decision on the price of the security. Therefore, both \( \sigma_t \) and \( \mu_t \) are small. At the opposite extreme, if \( V_t \) is very
large, then the trader begins to act more and more like an unconstrained trader. Since the trader is risk-neutral, the expected drift $\mu_t$ is pushed down to zero, 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 traders behave “as if” they were risk averse, and the equilibrium drift $\mu_t$ reflects this feature of the model. The expected return $\mu_t$ rises with $\sigma_t$.

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 2. We see that the Sharpe ratio is also single-peaked, rising and falling in roughly the same pattern with $\sigma_t$ and $\mu_t$.

### 3.2 Interpreting Equilibrium Volatility

We can offer the following intuition for the exact form taken by our ordinary differential equation (32), and for the multiplicity of the equilibrium. For the purpose of illustration, let us set the parameters so that $\alpha = \delta = 1$. Then, the equilibrium price satisfies

$$\ln P_t = rt + \eta z_t + \tilde{\sigma}_t V_t$$
where we denote by $\tilde{\sigma}_t$ the active trader’s conjecture about price volatility. From Itô’s Lemma,

$$\sigma_t = \eta \sigma_z + \tilde{\sigma}_t \times \text{(diffusion of } V_t) + V_t \times \text{(diffusion of } \tilde{\sigma}_t)$$

$$= \eta \sigma_z + V_t \left[ \tilde{\sigma}_t + V_t \frac{\partial \tilde{\sigma}}{\partial V_t} \right]$$

At equilibrium, $\sigma_t = \tilde{\sigma}_t$, which gives us the key ordinary differential equation (32).

Note the way that the active trader’s equity $V_t$ affects the solution. When $V_t = 0$, provided that the conjectured $\tilde{\sigma}$ is well-defined, at an equilibrium we have $\sigma_t = \eta \sigma_z$, the fundamental volatility. In this sense, the excess volatility above the fundamental volatility, given by $\sigma_t - \eta \sigma_z$ is the endogenous risk generated by the presence of active traders. It is the additional volatility that is generated due to the presence of traders who react to outcomes, and who in turn affect those outcomes themselves.

To see this, begin at the point where $V = 0$ and raise $V$ to a small positive number $\epsilon > 0$ under the assumption that the active traders conjecture that increased $V$ leads to an increase in conjectured volatility $\tilde{\sigma}$. Then, we must have an increase in actual $\sigma_t$. Under the conjectured increase in $\tilde{\sigma}$, we have

$$\sigma(\epsilon) = \eta \sigma_z + \text{positive term}$$

In other words, if the traders held the belief that $\sigma$ increases initially with $V$, then such beliefs would be confirmed in equilibrium.

The multiplicity of equilibrium associated with the arbitrary constant of integration $c$ can be understood in this context. The constant $c$ is a measure of the conjectured sensitivity between the state variable $V_t$ and the diffusion $\sigma_t$. If the active traders conjecture that $c$ is large (so that the feedback effects are assumed to be large), then they will act in such a way as to bring about the large feedback effects. The fact that the equilibrium condition results in a first-order ordinary differential equation means that there is some indeterminacy in tying down exactly how strong the feedback effects are from purely the consistency requirement that equilibrium reasoning imposes.

To summarize, the price-taking FIs’ market actions are expected, and at equilibrium confirmed, to influence the evolution of the price process themselves. While the intrinsic uncertainty (here the residual demand valuation
shocks) is the only randomness in the economy, it gets multiplied many fold by the presence of active traders, especially if they are undercapitalized. This we refer to as the phenomenon of *endogenous risk* (Danielsson and Shin (2003)). Risk is endogenous since FIs are not subjected to any fundamental preference or endowment shocks, and yet they add risk in a systematic, substantial, self-fulfilling and auto-exciting way.

The mechanism driving our one-factor model with a single state variable $V_t$ is the channel of endogenous risk appetite. In downturns, the VaR constraints bind harder, inducing feedbacks as asset sales beget asset sales, delevering begets delevering, and forcing the FIs to act in ever more risk-averse ways. Our model has the feature that once a crisis hits and risk-aversion and all the other factors peak, it will take time for risk-aversion to come down. This is borne out in the data as well (see Coudert et al (2008)). This is because the effect of risk-aversion on markets does not vanish after the uncertainty is resolved and the extent of the crisis becomes acknowledged. Our model predicts (consistent with the evidence) that risk aversion comes down over time as the financial cycle improves and the capital basis of FIs replenishes to more normal levels. There may then in fact obtain a long period of “moderation” with low volatility and high risk-appetite, such as the one that preceded the current crisis. During periods of moderation the economy is perceived to be far away from any serious pain points, as reflected in low Lagrange multipliers on risk constraints. The capital write-downs in 2007 and 2008 caused bank equity to shrink considerably, ending the “great moderation” and allowing the endogenous risk mechanism detailed above to hit the FI sector with full force.

While the single-factor version of our model performs well along some dimensions, there is no doubt that the world is more complex and has many factors. For instance, in our single factor model, we cannot accommodate history dependence on $V_t$. The outcome depends on the current value of $V_t$ only, and not on the path it took to reach the current level.

Allowing history dependence would enrich the dynamics considerably. History dependence may amplify the feedback effects considerably. Although such factors would not be needed to get sizable effects, there are certainly numerous other stylized facts that would need further factors to be explained satisfactorily. For instance the idea that the longer the period of high risk appetite, the larger the vulnerabilities that build up, and therefore the larger the resulting instabilities, would need a multi factor representation.
3.3 Volatility of Volatility

Major recent interest has been raised both in academia and by practitioners for the study of the volatility of volatility. There seems to be a consensus forming that the volatility of volatility (vol of vol) is itself random, rather than stable or mean-reverting over time, as well as being countercyclical (e.g. Pan (2002), Jones (2003), Corsi et al (2005)). In this simple one-factor model, vol of vol is generated by the dynamics of the capital process, which in turn depends on $\sigma_t$. So changing levels of volatility of the risky asset imply a high vol of vol of the risky asset.

Formally, vol of vol is the absolute value of the diffusion coefficient $\sigma_t^\sigma$ in $d\sigma_t = \mu_t \sigma_t dt + \sigma_t^\sigma dW_t$. It satisfies, by Itô’s Lemma, the following equation.

$$
\sigma_t^{\sigma} = \frac{V_t}{\alpha} \frac{\partial \sigma}{\partial V_t} = \frac{1}{\alpha V_t} \left[ \alpha^2 \delta(\sigma_t - \eta) - V_t \sigma_t \right]
$$

Figure 4: Volatility and $\sigma_t^{\sigma}$

Figure 4 plots both $\sigma_t$ and $\sigma_t^{\sigma}$. As we can see, our model generates a relationship that bears some resemblance to a countercyclical vol of vol. Whenever vol is increasing (either as equity is very small and growing, or very large and falling) vol becomes more volatile (the variance of vol goes up) and does so faster than vol itself. As equity further converges to that level at which vol is maximal, the variance of vol decreases again, indicating that the rate of change in vol diminishes as the maximal level of vol is getting closer.
This seems to be borne out in the data. Jones (2003) estimates a constant-elasticity of variance (CEV) generalization of the variance process in Heston (1993) and finds that the coefficient of elasticity is significantly larger than 0.5, indicating that on average vol of vol is higher when the level of vol is higher. This holds in our model for all levels of capital except those levels in a neighbourhood of the maximal vol level. Because Jones imposes an exogenous monotonic relationship between vol and vol of vol of the CEV type, the local non-monotonicity around the critical equity level that we find cannot be captured.

Similarly, when vol of vol is viewed as a function of capital (rather than as a function of vol as in Jones), vol of vol seems to predict market downturns two years hence. As capital levels are reduced, volatility increases but vol of vol increases and picks up the majority of its increment earlier than the spike in volatility itself (see Figure 4). By the time capital is further reduced to a level at which volatility is maximal (and is therefore not changing for small changes of equity), vol of vol needs to go to zero. This is because by Itô’s Lemma the vol of the vol is proportional to the slope of the vol curve, which is zero at the maximum.

This logic also implies that a one-factor model cannot generate vol of vol that is much higher than vol itself during a crisis, although it can generate a vol of vol that is a multiple of the vol during periods of sufficient capitalization.

---

4 Mele (2008), personal communication.
3.4 Derivatives Pricing Implications

It is well known that the Black-Scholes-Merton (Black and Scholes (1973), Merton (1973)) implied volatilities exhibit a negative skew in moneyness $K/S$ that is fading with longer time to maturity (see for instance Ait-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.

In our framework the REE volatility function $|\sigma_t|$ largely depends negatively on trading 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 – being the only one it can be 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.

Plot ?? shows the implied volatility smirk at a fixed maturity of 2 years, but with different initial capital levels. As expected, if $V(0)$ is very close to zero, then capital levels can only grow, so downside is limited and the upside bears a lot of volatility as FIs enter the market in force, and the smirk is upward sloping. This is a testable implication for developing economies or for periods of large financial deregulation. For all other initial capital levels, the smirk has the usual downward slope. The smirk is more pronounced the smaller the capital level is, as long as capital is beyond the critical level at which risk aversion plays havoc.

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) are counter-cyclical as well (again excluding FI capitals close to the zero-capital boundary). A healthier 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.”

Plot 6 gives the usual IV surface 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.

4 Equilibrium in the General Case

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

$$\frac{dP^i_t}{P^i_t} = \mu^i_t dt + \sigma^i_t dW_t$$  \hspace{1cm} (37)

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

$$dz^i_t = \sigma^i_t dW_t$$  \hspace{1cm} (38)
where $\sigma^i_z$ is a $1 \times N$ vector that governs which Brownian shocks affect the passive traders’ demands.

### 4.1 Solution for Case with $N$ Risky Assets

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

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

(39)

Also, we use the following shorthand.

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

(40)

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

(41)

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 (20) we have:

$$
\sigma^i_t = \tilde{\epsilon}^i_t V_t \tilde{\lambda}^\top_t \tilde{\sigma}^i_t + \eta \sigma^i_z
$$

(42)

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 $r 1_N$. 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 $(\tilde{\mu}, \tilde{\sigma})$ 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:

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

(43)
By stacking into a diffusion matrix, at a REE the diffusion matrix satisfies

\[ \sigma_t = V_t \epsilon_t \lambda_t^\top \sigma_t + \eta \sigma_z \]  \hspace{1cm} (44)

Using the fact that \( \lambda_t^\top \sigma_t \sigma_t^\top = \beta_t^\top \), \( \sigma_t \) 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

\[ (\sigma_t - \eta \sigma_z) \sigma_t^\top = V_t \epsilon_t \beta_t^\top \]  \hspace{1cm} (45)

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 (42) (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 feedback effect on vol
from VaR 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

\[ \epsilon_t := 1 - V_t \lambda_t^\top \epsilon_t \]

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

\[ \sigma_t = \eta \left[ \frac{V_t}{1 - V_t \lambda_t^\top \epsilon_t} \epsilon_t \lambda_t^\top + I \right] \sigma_z \]  \hspace{1cm} (46)

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

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

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

4.2 The Symmetric $N > 1$ Case

We examine a special case that allows us to solve for the equilibrium in closed form. The benefit of this is that we are able to reduce the dimensionality of the problem back to one 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 $\tilde{\sigma}_z I_N$ 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 a REE where $\mu^i_t = \mu_1^i$, $\beta^i_t = \beta_1^i$, $\lambda^i_1 = \lambda_1$, $\epsilon^i_t = \epsilon_1^i$, $\sigma_1^i = \sigma_1^{11}$ and $\sigma_1^{ij} = \sigma_1^{12}$, $i \neq j$. First, notice that $\epsilon_t \lambda_t^\top = \epsilon_1^\top \lambda_1^\top \mathbf{1}$, and that $\epsilon_t \lambda_t = N \epsilon_1^\top \lambda_1$, where $\mathbf{1}$ is a $N \times 1$ vector of ones (so that $\mathbf{1} \mathbf{1}^\top$ is the $N \times N$ matrix with the number 1 everywhere).

From (46) we see that the diffusion matrix is given by

$$\eta \sigma_z \left( \frac{V_t \lambda_1^\top \epsilon_1^i}{1 - NV_t \lambda_1^\top \epsilon_1^i \mathbf{1} \mathbf{1}^\top + I} \right)$$  \ (48)

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 $\sigma_t^{ij}$ 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 $\eta$ replaced by $\frac{\eta}{N}$, i.e. $x_t$ is equal to the right-hand-side of (33) 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 $\sigma_{i}^{ii} = x_t + \frac{N-1}{N} \eta \tilde{\sigma}_z$, and for $i \neq j$, $\sigma_{i}^{ij} = x_t - \frac{1}{N} \eta \tilde{\sigma}_z$. Also, $\Sigma_{i}^{ii} = \text{Var}(\text{return on security } i) = \eta^2 \tilde{\sigma}_z + \frac{1}{N} (N^2 x_t^2 - \eta^2 \tilde{\sigma}_z^2)$, and for $i \neq j$, $\Sigma_{i}^{ij} = \text{Cov}(\text{return on security } i, \text{return on security } j) = \frac{1}{N} (N^2 x_t^2 - \eta^2 \tilde{\sigma}_z^2)$ and $\text{Corr}(\text{return on security } i, \text{return on security } j) = \frac{N x_t^2 - \eta^2 \tilde{\sigma}_z^2}{N x_t^2 + \eta^2 \tilde{\sigma}_z^2}$.

Risky holdings are $D_t = \frac{V_t}{\alpha N^{3/2} x_t}$.

The risk-reward relationship is given by

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

(49)

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. On the one hand this means that less than the full impact of the shock on security $i$ will be transmitted into the security return $i$, potentially leading to a less volatile return. The reason is that a smaller fraction of the asset portfolio is invested in security $i$, reducing the extent of the feedback effect. On the other hand, the demand shocks to securities 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 $\sigma_{i}^{ij} = x_t - \frac{1}{N} \eta \tilde{\sigma}_z$, 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.

For comparison purposes, denote the scalar diffusion coefficient from the $N = 1$ case, as given by (33), by $\sigma_{i}^{N=1}$, and fix $c = 0$. The first direct
effect can be characterized as follows: $\sigma^{N=1}_t < \sigma^{N=1}_t$ if $\eta \tilde{\sigma}_z < \sigma^{N=1}_t$. 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 $V^* : \sigma^{N=1}_t(V^*) = \eta \tilde{\sigma}_z$, the direct effect is larger than in the $N = 1$ economy\footnote{For instance, as $V \to \infty$, we have $\lim_{V \to \infty} \sigma^{N=1}_t = \frac{N-1}{N} \eta \tilde{\sigma}_z > 0 = \lim_{V \to \infty} \sigma^{N=1}_t$.} 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 ($\lim_{V \to \infty} \sigma^{N=1}_t = 0$), with $N > 1$ on the other hand individual volatility remains ($\lim_{V \to \infty} \Sigma^{N=1}_i = \frac{N-1}{N} \eta^2 \tilde{\sigma}^2_z > 0$) 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-security case if $V$ is small: $\Sigma^{N=1}_i (\sigma^{N=1}_t)^2$ if $\eta^2 \tilde{\sigma}^2_z / N^2 < x^2_t$. 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 forcefully for small levels of capital, the fact that idiosyncratic shocks are mixed and affect all securities implies that security 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 $V$ for the $c = 0$ case.

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

Figure 8: $N = 2$, $\Sigma^{ii}$ and $\Sigma^{ij}/\Sigma^{ii}$
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 Concluding Remarks

We have examined a rational expectations model of stochastic volatility with the feature that traders act as if their preferences are changing in response to market outcomes. In this sense, we have shown how risk appetite and risk are determined together and how both are tied to market outcomes. The channel through which such apparent preferences and beliefs are linked are the risk constraints. As risk constraints bind harder, effective risk aversion of the traders also increases. We have argued that this simple story of risk aversion feedbacks can explain much of the observed counter-cyclical features of volatility, vol of vol, correlations, implied vols, risk premia and Sharpe ratios. They can all be thought of as being driven by the single factor of endogenous risk aversion.

Our discussion has focused purely 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 2007-8. Brunnermeier et al. (2009), Danielsson and Zigrand (2008) and Morris and Shin (2008) are recent discussions on how the debate on the future of financial regulation can take account of the themes explored here.
Appendix

Lemma 3 [Properties of the diffusion term] [S1] \( \lim_{t \to 0} \sigma(V_t) = \eta \sigma_z \)
[S2] \( \lim_{t \to \infty} \sigma(V_t) = 0 \) and \( \lim_{t \to \infty} V_t \sigma(V_t) = \infty \)
[S3] \( \lim_{t \to 0} \frac{\partial \sigma}{\partial t} = \frac{\sigma_t}{\sigma^2} \) and \( \lim_{t \to 0} \frac{\partial^2 \sigma}{\partial^2 t} = \frac{4 \eta \sigma_z}{(\alpha \sigma)^2} \). [Call \( f(V) := \frac{\sigma - \eta \sigma_z}{V} \)]
and notice that \( \lim_{V \to 0} f(V) = \lim_{V \to 0} \frac{\sigma}{V} \). Since we know the expression for \( \frac{\partial \sigma}{\partial t} \) by (32), we see that the problem can be transformed into \( \lim f = \lim \frac{1}{V} (\alpha^2 \delta f(V) - \eta \sigma_z) \). In turn, we can replace \( \frac{\partial}{\partial V} \) definitionally by \( f + \frac{\eta \sigma_z}{V} \) to get to \( \lim f = \lim f(V) [\alpha^2 \delta f(V) - \eta \sigma_z] \). 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 \( V^* \), \( \sigma \) is strictly decreasing.
[The second observation comes from (32) while the first one comes from the fact that the mapping \( V \mapsto \int_{-\infty}^{\infty} \frac{e^{-u}}{V} du \) is a bijection between \( \mathbb{R}_+ \) and \( \mathbb{R} \), so for each chosen \( c \), there is a unique \( V(c) \) setting \( [\frac{c - \alpha^2 \delta \eta \sigma_z}{V(c)} \frac{e^{-u}}{V} du] = 0 \).
[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 (48) the variables of interest as \( \sigma_t^1 = \eta \tilde{\sigma}_z \left( \frac{\tilde{V}_t \tilde{\lambda}_t}{1 - N \tilde{V}_t \tilde{\lambda}_t} + 1 \right) \) and \( \sigma_t^2 = \eta \tilde{\sigma}_z \left( \frac{\tilde{V}_t \tilde{\lambda}_t}{1 - N \tilde{V}_t \tilde{\lambda}_t} \right) = \sigma_t^{11} - \eta \tilde{\sigma}_z \).

Next, we compute the variance-covariance matrix, the square of the diffusion matrix (48):
\[
\Sigma_t = \sigma_t \sigma_t = \eta^2 \tilde{\sigma}_z^2 [I_N + m_t \mathbf{1}\mathbf{1}^T] = \eta^2 \tilde{\sigma}_z^2 I_N + g_t \mathbf{1}\mathbf{1}^T
\]
where
\[
m_t := N \left( \frac{\tilde{V}_t \tilde{\lambda}_t}{1 - N \tilde{V}_t \tilde{\lambda}_t} \right)^2 + 2 \frac{\tilde{V}_t \tilde{\lambda}_t^2}{1 - N \tilde{V}_t \tilde{\lambda}_t^2}
= \frac{1}{\eta^2 \tilde{\sigma}_z^2} (\sigma_t^{11} - \eta \tilde{\sigma}_z) [2 \eta \tilde{\sigma}_z + N (\sigma_t^{11} - \eta \tilde{\sigma}_z)]
\]
\[
g_t := m_t \eta^2 \tilde{\sigma}_z^2
\]
where we used the fact that \( \eta \tilde{\sigma}_z \left( \frac{\tilde{V}_t \tilde{\lambda}_t}{1 - N \tilde{V}_t \tilde{\lambda}_t} \right) = \sigma_t^{11} - \eta \tilde{\sigma}_z \). Then insert \( \Sigma_t \) into the reward-to-risk equation \( \Sigma_t \lambda_t \mathbf{1} = \frac{\mu_t - r}{\sqrt{\lambda_t}} \mathbf{1} \) to get \( \sqrt{\xi_t \lambda_t} \left[ \eta \tilde{\sigma}_z + N (\sigma_t^{11} - \eta \tilde{\sigma}_z) \right]^2 = \mu_t - r .
\]
Next compute $\xi_t$. By definition, $\xi_t := (\mu_t^1 - r)^2\Sigma_t^{-1}\mathbf{1}$. Since $1 + g_t(\eta\tilde{\sigma}_z)^{-2}N \neq 0$, by the Sherman-Morrison theorem we see that

$$\Sigma_t^{-1} = (\eta\tilde{\sigma}_z)^{-2}I - \frac{g_t}{(\eta\tilde{\sigma}_z)^4 + N(\eta\tilde{\sigma}_z)^2g_t}\mathbf{11}^\top$$

and therefore that

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

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

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

where $A := \mu_t^1 - r$ 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

$$\beta_t^1 = \frac{t_{AB}}{\sqrt{N}}[\eta\tilde{\sigma}_z + N(\sigma_t^{11} - \eta\tilde{\sigma}_z)]$$

By definition of $\epsilon_t^1$:

$$\epsilon_t^1 = \frac{1}{\sigma^2}\frac{t_{AB}}{\sqrt{N}}\left[\frac{\sqrt{N}[\eta\tilde{\sigma}_z + N(\sigma_t^{11} - \eta\tilde{\sigma}_z)] + V_t\sqrt{N}\frac{\partial\sigma_t^{11}}{\partial V_t}}{1 - \lambda_t^1}\right]$$

(50)

Inserting all these expressions into the equation for $\sigma_t^{11}$, $\sigma_t^{11} = \eta\tilde{\sigma}_z\frac{1 - (N - 1)V_t\lambda_t^1}{1 - N\lambda_t^1}$ 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 $\eta$ replaced by $\eta/N$ and where $\sigma(V)$ is replaced by $x(V)$.

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

Finally we compute the risk premia. Using Itô’s Lemma on (20) we get

$$\mu_t^1 - \frac{1}{2}(\sigma_t^{11})^2 = \alpha\epsilon_t^1(\text{drift of } V_t) + rdt + \frac{1}{2}\alpha\frac{\partial\epsilon_t^1}{\partial V_t}(\text{diffusion of } V_t)^2$$

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

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

33
and using $\sigma_t = (x_t - \frac{1}{N} \eta \tilde{\sigma} z) 11^T + \eta \tilde{\sigma} z I$ the squared diffusion term can be verified to be equal to $\frac{V_t^2}{\alpha^2}$. The drift equation becomes

$$
(\mu_1^T - r) \left[ 1 - \epsilon_1^T \frac{V_t}{\sqrt{N}} x_t \right] = \frac{1}{2} (\sigma_t^{11})^2 + \alpha^2 \epsilon_1^T r V_t + \frac{1}{2} \frac{V_t^2}{\alpha} \frac{\partial \epsilon_1^T}{\partial V_t}
$$

We can rewrite (5) by inserting the ODE for $x_t$ to get rid of the partial derivative term:

$$
\epsilon_1^T = \iota_A \iota_B \frac{\sqrt{N}}{V_t} \left( x_t - \frac{\eta}{\sqrt{N}} \tilde{\sigma} z \right)
$$

Performing the differentiation of $\epsilon_1^T$ and inserting into the drift equation completes the proof.

References


