No News is Good News

A thesis presented
by
Wei Sun
to
The Committee on Degrees in Economics
in partial fulfillment of the requirements
for the degree of
Bachelor
in the subject of
Economics

Williams College
Williamstown, Massachusetts
May 2013
Abstract
This paper explores investor behavior in an environment of financial innovation. It develops a model in which investors are given a choice between a risk-free asset and one that promises a higher current yield, but includes a nonzero but unknown probability of a large loss in value. Lacking any other information about the asset’s fundamentals, investors use past performance to make inferences about the collapse probability: the more time that passes without a collapse, the safer it is thought to be. The model generates behavior that is suggestive of herding or bubbles, but without appealing to agency problems, information asymmetries, or deviations from rational expectations. The paper also considers the effect of monetary policy, namely lowering risk-free interest rate, on risk taking in this framework.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>1 Introduction and Background</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Literature Review</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Overview of the model</td>
<td>5</td>
</tr>
<tr>
<td>2 The Portfolio Allocation Decision</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Risk-neutral case</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Risk-averse case</td>
<td>14</td>
</tr>
<tr>
<td>2.3 Connection to behavioral finance</td>
<td>19</td>
</tr>
<tr>
<td>3 General Equilibrium and Empirical Implications</td>
<td>20</td>
</tr>
<tr>
<td>3.1 Incorporating price</td>
<td>20</td>
</tr>
<tr>
<td>3.2 Empirical Implications</td>
<td>27</td>
</tr>
<tr>
<td>4 Conclusion</td>
<td>30</td>
</tr>
<tr>
<td>A Theorems and Proofs</td>
<td>32</td>
</tr>
<tr>
<td>A.1 A Note on Computation</td>
<td>32</td>
</tr>
<tr>
<td>A.2 Proof of Lemma A.1.4</td>
<td>34</td>
</tr>
<tr>
<td>References</td>
<td>36</td>
</tr>
</tbody>
</table>
Acknowledgments

I would like to thank a number of people who have made this thesis possible—the past year was especially tough for me, and I thank you all for always having confidence in me, and encouraging me to keep going.

To start with, I am obliged to Professor Kenneth Kuttner for his guidance throughout the year and his patient readings of every draft. His wisdom also made the concurrent and tormenting process of graduate school application and decision much more pleasant. I have learned from him much more than economics theories alone. I am indebted to Professor Gerard Caprio, Professor William Gentry, Professor David Love, Professor Peter Pedroni, and Professor Michael Rolleigh not only for their excellent suggestions on how to revise my thesis, but also on kindling my interest in Economics. Their wonderful classes have provided many inspirations, and have taught me what is “good economics”. I would like to thank Professor Bakija for his excellent feedback during thesis meetings, and his constant email reminders to keep me the procrastinator on track (more or less). I thank the Economics Department at Williams College for generous financial support through the Carl Van Duyne Prize, and for providing me with challenges and precious opportunities during my undergraduate years—those have exceeded my wildest dreams.

I would also like to thank my friends in Williams, in Cambridge, back at home and in other places—because of you, I am a happy person and that made me creative. In particular, I would like to thank Sungik Yang and Qiao Zhang for regular exchange of ideas over my undergraduate years. Last but not least, I would like to thank my parents for never giving up on their naughty son in the past twenty-two years. I am extremely lucky to be your son.

I welcome any comment on the paper. All errors are my own.

Contact: sun.wei.199007@gmail.com
Chapter 1

Introduction and Background

1.1 Introduction

Recent experience has shown that understanding the boom and bust cycle of financial market is of first-order impotance for macro stability. This paper explores agents’ unfamiliarity with the asset as a source of instability of financial markets. The idea is the following: when the agents do not have complete information about an asset, they estimate parameters of the asset (like return and variance) using past history, which themselves are observed random variables.\(^1\) Since those observables are random, there is a non-trivial probability that they will mislead the investor temporarily,\(^2\) when beliefs are later updated (closer towards truth) the price exhibits something very similar to that of a boom and burst cycle.

Unfamiliarity is at the heart of our model, and this is motivated by empirical research on the topic. Kindleberger and Aliber (2011) note that “events that lead to a crisis start with a 'displacement', some exogenous, outside shock to the macroeconomy system. Reinhart and Rogoff (2009) echo this observation and coined the phrase “This-Time-Is-Different

\(^1\)In other words, if we follow the Bayesians and treat the parameter of interest as random variable, the mutual information of those observations and that of the parameter of interest is strictly positive.

\(^2\)Just like for a fair coin toss, there is always a positive probability that one will get a sequence of all heads, and once the coin is tossed large enough times, the probability one will observe a sequence of all heads goes to one. In financial market, when there are enough times when people collect (random) information, the probability they will be seriously mislead at least once goes to one. We have observed many boom and burst cycles, but there could be infinitely many more boom and burst cycles.
syndrome” to lament that agents too readily believe that a new paradigm has arrived because of perceived “sound fundamentals, structural reforms, technological innovation, and good policy”. For example, the improved Britain’s economy in mid-1840’s led to the railway mania. The technological innovations in the IT industry generated much excitement in the stock market which ultimately ended with a crash. The theme is that the agents are in a new economy when old information is no longer relevant. It is this unfamiliarity that creates room for instability. Furthermore, it is also an empirical regularity that boom and bust cycle often follows financial innovation and liberalization (Kaminsky and Reinhart (1999)). For example, the South Sea Bubble followed an innovative idea John Blunt had about handling the government debt. The financial liberalization in Japan and the surge in the foreign exchange value of yen in the 1980s was followed by the boom and bust in the real estate and stocks. The recent financial crisis followed the financial innovations such as Credit Default Swap (DCS) and other exotic financial derivatives. It is following financial innovation and liberalization that many new and unfamiliar assets appear, and when a new era has begun that old information is believed to be less relevant. This paper focuses on an unfamiliar asset, about which the agents have no complete information and rely on observables to make inferences and decisions. An asset could be unfamiliar for many reasons. For example, the bond of a company could be unfamiliar, even if the company might have issued the bonds for many decades, as long as the agent believes that he is in a new paradigm where the past information is not very relevant for our current optimization problem. However, an asset could be entirely new, but not unfamiliar—for example, if the new asset is based on an underlying asset that we know very well, then via calculation, this new asset could be as familiar as the old asset. However, pricing of some multi-asset derivatives could be very sensitive to the assumed covariance structure of the underlying assets. Since the underlying covariance can only be estimated with historical data, the information we have about the asset is far from sufficient.

The purpose of this paper is three fold. The first is to characterize the behavior of the investor when faced with an unfamiliar asset, namely his demand of the unfamiliar asset

---

3In reality, the agent never receives unambiguous information that a change of paradigm has taken place, this could also be a decision—discarding the previous prior on $\lambda$ for the new prior $\pi(\lambda)$. Ortoleva (2012) modeled this process as Hypothesis Testing.
Chapter 1: Introduction and Background

changes as time passes by and to monetary policy shock. The second aim is to characterize
the price of the unfamiliar asset in a general equilibrium setting. Finally this paper aims to
build a model that could incorporate behavioral elements in a tractable way.

In this paper, Section 1.2 first review the relevant literature and how this paper relates
to it. Section 1.3 gives an overview of the model and explains the intuition of the model.
Chapter 2 studies the partial equilibrium where the price (premium) is exogenous. Section
3.1 endogenizes the premium and considers the general equilibrium of the financial market.
Section 2.3 is a digression on how the model can incorporate behavioral elements. Section
3.2 discusses the empirical implications of the model. Finally, Section A.1 discusses some
of the computational issues related to the model.

1.2 Literature Review

The theoretical literature has taken five different approaches in explaining the source of
instability of financial market. The first school of thought has ascribed the boom and crash
to changes in fundamentals (Garber (1990, 2001)). A second school points to speculation
among traders without a common prior belief (Harrison and Kreps (1978)), which in partic-
ular includes the “greater fool” models (Allen and Gorton (1993); Conlon (2004)).

A third school points to agency problems like moral hazard or risk-shifting, where fund managers’
incentives are distorted because they do not bear all of the losses if the investment turns
out to be unprofitable (Allen and Gale (2000); Allen and Gorton (1993); Barlevy (2007)).
Fourth, researchers have documented behavioral irrationality as an explanation. In this
case, rational arbitragers were prevented from fully correcting the prices by fundamen-
tal risk due to lack of perfect arbitrage; noise trader risk (DeLong et al. (1989); Shleifer
and Summers (1990); Shleifer and Vishny (1997)); and synchronization risk (Abreu and
Brunnermeier (2003)). The behavioral finance literature has supplemented this school with
experimental evidence on systematic biases when people form beliefs such as overconfi-
dence (Alpert and Raiffa (1982), Fischhoff et al. (1977)), optimism and wishful thinking
(Buehler et al. (1994)), and under-reaction (with small quantities of similar information)

\footnote{The initial prior does not have to be different as shown by (Allen and Gorton (1993)).}
and over-reaction (after large quantities of similar information) (Barberis et al. (1998)). Finally, a fifth strand of literature incorporates learning to explain the excess high risk premium of stock return, and the excess volatility in addition to the financial instability, and my works falls into this category.

For example, in Timmermann (1993) the agent knows the structure of the dividend process (geometric Brownian Motion) of a stock, but does not know the parameters characterizing the process. The agent then uses optimal (frequentist) formula to estimate those parameters and price the assets using those estimates. This model was able to explain predictability of stock prices and excess volatility. In Veronesi (2004), there are two states of the world—good state and bad state, and dividend is expected to be lower in bad state. However, the agent does not know which state he is in, and uses dividend information to make inference about which state the economy is in. This model explains high risk premia, time-varying volatility, asymmetric reaction to good and bad news, and excess return volatility. This paper is in this tradition. In one sense, we are following the “bounded rationality” tradition, which has proposed to “create theories with behavioral foundations by eliminating the asymmetry the rational expectaions builds in between the agents in the model and the econometrician who is estimating it”, that is “requiring that the agents in the model be more like the econometricians” (Sargent (2011)). This paper lets the agents be “Bayesian econometricians”, updating their prior with relevant information and optimizing accordingly. However, this paper differs from the papers in this school in several different aspects. First, the asset in this paper is motivated by fixed-income assets like CDOs, whose surprising and large negative returns were a major player in this financial crisis, rather than equity. Second, this paper does not involve different states, and the agent is only making inference about a constant parameter. Finally, the basic version of the model assumes reationality, but I will show in Section 2.3 how it can generalized in include behavioral elements: indeed with suitable data the model could test for the existence of irrationality, in particular under-reaction and over-reaction.
1.3 Overview of the model

Investors choose between two assets. One is the standard risk-free asset, which pays a known return of $r_0$. The other is a risky asset. Rather than modeling the return in terms of mean and variance, as in Markowitz (1952), this paper is interested in studying those rare but large negative realized payoffs. Thus in the model, the risky asset resembles a bond. Specifically, the asset pays a return in excess of the risk-free rate for some extended period of time (until the stopping time $T$), after which it “blows up” or crashes and the investor experiences a large, negative return. This is intended to mimic the behavior of many of the assets such as Collateralized Debt Obligations (CDOs) and Credit Default Swaps (CDSs), which paid positive excess returns for periods measured in years, and subsequently crashed. When the underlying blow-up frequency is sufficiently low, this is also very reminiscent of the “peso problem” where there is a return differential between two seemingly identical risk-free assets. Indeed our model will deliver such return differential except that that differential will shrink overtime.

Specifically, I assume that the return follows the following distribution:

$$\text{payoff} = \begin{cases} r_0 + \alpha & \text{when there is no crash} \\ -L & \text{when the asset crashes} \\ 0 & \text{after the asset crashes} \end{cases} \quad (1.1)$$

The $\alpha$ parameter can be interpreted as a risk premium. For now I will assume it is exogenous. Section 3.1 will make it endogenous.

Time is continuous. The probability of a crash in time period $(t, t+\Delta t)$ is $\lambda \cdot \Delta t + o(\Delta t^2)$. This implies that the expected duration of a non-crash period is $\frac{1}{\lambda}$, which in turn implies that the expected payoff of the asset until blow up is $\frac{1}{\lambda} (r_0 + \alpha) - L$. More succinctly, the blow up disaster is modeled as a Poisson Process with intensity parameter $\lambda$.

A departure from standard asset pricing theory is that the probability distribution of the asset return (parameterized by $\lambda$) is not known to the investor. Consequently, he has to make inferences about the unknown $\lambda$ based on its track record (i.e. how long is has gone without a crash). The natural way to do this is to have the investor use Bayes’ rule to update his prior of the $\lambda$. 

The investor’s decision making process proceeds as follows. At time 0, he chooses the optimal allocation of his wealth (normalized to 1) between the two assets. After $\Delta t$, if the asset does not crash, he receives the return of $[p(r_0 + \alpha) + (1-p)r_0]\Delta t$ where $p$ is the proportion of his wealth alloted to the unfamiliar asset. He then uses this information to calculate the posterior on $\lambda$. In continuous time, $\Delta \to 0$, and he optimizes his portfolio continuously.

If the asset does crash, the return of $-L$ is realized, and the process stops. The investor calculates the average payoff $r - \frac{L}{T}$, and reinvests his original principal. I follow the convention in using the von Neumann-Morgenstern expected utility framework to model the investor’s decision, rather than a “behavioral” alternative such as prospect theory. Specifically, I assume that the investor is maximizing the expected utility derived from the asset’s average return $r - \frac{L}{T}$. The model makes the simplifying assumption that the agent only consumes out of his return on investment, and invests the fixed amount he is endowed with initially, regardless of contingencies.

**Example** (Numerical example). A risk-free asset pays off $r_0 = 0.05$ in each period and it costs 1 dollar for each unit of asset. The unfamiliar asset will deliver $r = 0.08$ in each period until it blows up. The agent is endowed with one hundred dollars, and he purchases a hundred shares of the unfamiliar asset. When the asset blows up, the agent will stop receiving any payoff from the asset (equivalent to possessing no asset) and does not recover the one hundred dollars he paid for the asset, that is his loss will be $100L = 100$. For example, the asset keeps performing until 100 periods later, which gives the agent a cumulative return of $800$, which after subtracting the loss of one hundred dollars, is a net $700$ gain. The average consumption in each period is thus $7$ per period. At the end of 100 periods, the agent will re-invest his $100$ in a combination of risk-free and risky assets.

Chapter 2 studies the investor’s decision in a partial equilibrium setting, taking the risk premium $\alpha$ as exogenous. Specifically, we study how the optimal allocation changes over time. Section 3.1 endogenizes risk premium and studies how prices of this asset change over time.
Chapter 2

The Portfolio Allocation Decision

This chapter works through the investors’ problem of optimal asset allocation. It begins in section 2.1 with the risk-neutral special case, i.e. a linear utility function. I am able to show that for each investor with prior $\pi(\lambda)$ there is a a threshold $t^*$—if and only if the asset does not blow up until $t^*$, the investor will completely switch to the risky asset from $t^*$ on. Section 2.2 considers the more complicated risk-averse case. It turns out that for a risk-averse investor with prior $\pi(\lambda)$, he will move to allocating positive proportion of his wealth to the risky assets.

2.1 Risk-neutral case

In this section, we study the case of a risk-neutral investor, which is arguably an appropriate assumption for some institutional investors. We will show that for any risk-neutral investor with a non-dogmatic prior of the blow-up frequency $\lambda$, and given an exogenous $\alpha$, as long as the asset does not blow up for long enough time (a long enough track record), the investor will find it optimal to switch to the asset and remain fully invested (existence of threshold). To show this, we first formalize the updating procedure of the investor, then we characterize conditions where investing in the risky asset is optimal. We propose two sufficient conditions for the existence of threshold and prove that imposing a non-dogmatic prior will satisfy these two conditions.

We now discuss the updating process of the investor. The investor begins with a prior
distribution over $\lambda$, denoted as $\pi(\lambda)$. “No news is good news”, in the sense that if $t$ passes without a crash, it leads the investor to revise downward his prior on the value of $\lambda$ (formalized as a chain of stochastically dominated distributions). This is a straightforward application of Bayes’ rule:

$$\pi(\lambda | N_t = 0) = \frac{\pi(\lambda) e^{-\lambda t}}{\int \pi(\lambda) e^{-\lambda t} d\lambda}, \tag{2.1}$$

where $N_t = 0$ denotes that the number of crashes is zero at time $t$.

For simplicity we will write $N_t$ instead of $N_t = 0$. The probability of no crash (conditional on $\lambda_0$) is $e^{-\lambda_0 t}$, and multiplied by $\pi(\lambda_0)$ (the numerator) is the joint probability of having $\lambda = \lambda_0$ and having no crash until time $t$. The marginal distribution for the waiting time $T|N_t$ has the following PDF:

$$f_t(y) = \int_\lambda \lambda e^{\lambda y} \pi(\lambda | N_t). \tag{2.2}$$

The alternative for investment in this unfamiliar assets is to invest in a risk-free asset, which gives a return $r_0$ per period per dollar invested. The expected excess return of an investment in the unfamiliar asset relative to an investment in the risk free asset is

$$E(rT - L | N_t) = (r - r_0)E(T | N_t) - L.$$

The investor will invest in the risky asset if and only if

$$E(T | N_t) \geq \frac{L}{r - r_0}.$$

In the rest of this section, we will prove that for an arbitrary pair of $(\alpha, \pi(\lambda))$, there exists a threshold $t^*$ in the sense that:

- The investor will not invest in the risky asset before time $t^*$.
- If the risky asset has not blown up at time $t^*$, the investor will switch to this risky asset and remain fully invested in it (until blow-up).

To prove the claim above, we only need to prove

$$\lim_{t \to \infty} E(T | N_t) > \frac{L}{r - r_0}, \quad \text{and} \quad \frac{\partial E(T | N_t)}{\partial t} > 0. \tag{2.3}$$

\footnote{By our assumption that the stochastic process ends when the blow-up takes place, $N_t = 0$ or 1.}
Equation (2.3) means that if the asset does not blow up for long enough, the investor’s posterior expectation of non-crash duration will be sufficiently large to justify investing in this risky asset. Equation (2.4) means that the posterior expectation will increase monotonically, and hence no investment before \( t^* \) and remain fully invested after \( t^* \). The above two conditions are sufficient conditions for the existence of threshold.

The following numerical example illustrates how the investor forms a posterior belief about \( \lambda \) and \( T \), and makes decisions accordingly.

**Example (Numerical example continued).** An individual has a prior of the blow-up intensity \( \lambda \) of this unfamiliar asset

\[
\pi(\lambda) = 4x \exp(-2x)
\]

or \( \frac{1}{2} \) Gamma(2) distribution.\(^2\) The prior on \( T \) is then \( F(2,4) \)

If at the end of the fourth period, the asset has still not blown up, then it is equivalent to observing \( \text{Poisson}(4\lambda) = 0 \), then the posterior on \( \lambda \) becomes

\[
\pi_4(\lambda) = 36\lambda e^{-6\lambda}
\]

or \( \frac{1}{6} \) Gamma(2).\(^3\)

The posterior distribution on \( T \) then becomes an Inverse-Gamma mixture of Exponential distribution.

\[
f_4(T) = \frac{72}{(6+T)^3}
\]

or \( 3F(2,4) \) distribution.

The expected net payoff (from one unit of the risky asset) until the blow-up is

\[
0.08 \times E(T | N_t) - 1 = -0.52
\]

Since the net expected payoff is negative, it is not optimal to invest in the unfamiliar asset.

We now prove equation (2.4). As remarked, this guarantees that the posterior expected payoff of the risky asset monotonically increases as time passes (assuming no crash), and

\(^2\)To avoid confusion between two conventions of scale vs. rate parameters, we put the scale parameter in front of the distribution, which is more intuitive.

\(^3\)That the posterior on \( T \) still follows a Gamma distribution is not coincidence, but rather the result of our carefully chosen prior, which will be discussed in further detail in section A.1.
Figure 2.1: Prior and Posterior Distribution of $\lambda$ and $T$
the optimal allocation of the risky asset will jump from 0 to 1 at a threshold time and remain there. We start with the monotonicity of posterior expectation on the intensity parameter.

**Theorem 2.1.1.** With the distribution specified in (2.1), we have

\[
\frac{\partial E(\lambda^{-1}|N_t = 0)}{\partial t} > 0
\]

For notational convenience, we use \( E_t(\cdot) \) to denote \( E(\cdot|N_t = 0) \).

**Proof.** We check that the conditions for differentiation under integration are satisfied.

\[
\frac{\partial E_t(\lambda^{-1})}{\partial t} = \frac{\partial}{\partial t} \int \frac{\pi(\lambda)e^{-\lambda t}}{\int \pi(\lambda)e^{-\lambda t} d\lambda} \lambda^{-1} d\lambda
\]

\[
= \int \frac{\pi(\lambda)e^{-\lambda t}(-\lambda)(\int \pi(\lambda)e^{-\lambda t} d\lambda) - \pi(\lambda)e^{-\lambda t}(-\lambda)\int \pi(\lambda)e^{-\lambda t}(-\lambda) d\lambda}{(\int \pi(\lambda)e^{-\lambda t} d\lambda)^2} \lambda^{-1} d\lambda
\]

\[
= -1 + E_t(\lambda)E(\lambda^{-1})
\]

\[
= E_t(\lambda)E(\lambda^{-1}) - E(\lambda \cdot \lambda^{-1})
\]

\[
= -\text{cov}(\lambda, \lambda^{-1}|N_t) > 0
\]

**Remark.** This is a weaker version of Theorem 2.2.1, which states that

\[
t_1 > t_2 \iff T|N_{t_1} \succ T|N_{t_2}
\]

where \( \succ \) stands for “First-Order stochastic dominates”.

**Corollary 2.1.2.** Condition (2.4) is satisfied.

**Proof.** This is a consequence of the Law of Iterated Expectation.

We now move on to Equation (2.3), the condition that ensures that the investor will eventually purchase the risky asset. However, this condition is not always satisfied. Consider the following counter-example

\(^4\)X₁ is said to first order stochastically \( X₂ \) if \( P(X₁ \leq c) \leq P(X₂ \leq c) \) for all \( c \) and the inequality is strict for some \( c \).
**Example.** Consider a dogmatic prior where an individual puts zero probability on low \( \lambda \):

\[
\pi(\lambda) = \begin{cases} 
0, & \lambda \leq \bar{\lambda} \\
> 0, & \lambda > \bar{\lambda}
\end{cases}
\]

In this case, despite observing no disaster for a long time (even for an infinite amount of time), in the posterior only \( \lambda > \bar{\lambda} \) will be given a positive probability. If \( \bar{\lambda} > \frac{r}{L} \), then we know that

\[
E(T|N_t) < \frac{1}{\bar{\lambda}} < \frac{r}{L}, \forall t
\]

that is, the expected stopping time \( T \) will never be long enough to make it optimal for the investor to buy the risky asset.

The idea is that when the investor has a dogmatic prior, regardless of what he observes, he fails to adequately update his posterior: Even if the asset has not blown up for an extended period of time, making a high \( \lambda \) extremely unlikely, the investor still assigns zero probability to lower values of \( \lambda \) and probability of 1 to higher values of \( \lambda \). If we rule out such dogmatic prior, condition (2.3) will be satisfied. It is worthwhile then to define what is meant by a non-dogmatic prior.

**Definition 2.1.1.** For a parameter \( \theta \) in the space of \( \Theta \), a prior \( \pi(\theta) \) is said to be non-dogmatic, if for any Lebesgue-measurable set \( A \) with positive Lebesgue measure, we always have

\[
\int_A \pi(\theta) > 0
\]

**Theorem 2.1.3.** Given a non-dogmatic prior, (2.3) is satisfied.

We first prove a useful lemma.

**Lemma 2.1.4.** For a non-dogmatic prior \( \pi(\lambda) \), for any \( \lambda^* > \inf \Theta \), if there is no crash, the tail probability that \( \lambda > \lambda^* \) goes to zero.

\[
z_t := P(\lambda \geq \lambda^*) = \int_{\lambda \geq \lambda^*} \frac{\pi(\lambda)e^{-\lambda t}}{\pi(\lambda)e^{-\lambda t}d\lambda} \to 0
\]
Proof. Let $q_t$ denote
$$q_t = \int_{\lambda \leq 0.5\lambda^*} \frac{\pi(\lambda)e^{-\lambda t}}{\int \pi(\lambda)e^{-\lambda t} d\lambda}$$
We prove that $z_t \to 0$.

Because $\pi(\lambda)$ is non-dogmatic, $q_0 > 1$.

$$z_t < \frac{z_0 \cdot e^{-\lambda^* t}}{q_0 \cdot e^{-0.5\lambda^* t} + (1 - z_0 - q_0)e^{-\lambda^* t}} = \frac{z_0}{q_0 \cdot e^{0.5\lambda^* t} + (1 - z_0 - q_0)} \to 0$$

We are now ready to prove Theorem 2.1.3 using Lemma 2.1.4 to construct lower bounds.

Proof. For any positive real number $c$, denote $\lambda^*$ such that
$$\frac{1}{\lambda^*} > c$$
$$E_t(\lambda^{-1}) \geq (1 - p_t)\frac{1}{\lambda^*} > 2c(1 - p_t) \to c$$
That is
$$\lim E_t(\lambda^{-1}) \to \infty$$
and this is a sufficient but not necessary condition for (2.3).

We conclude this section with a theorem characterizing the threshold, and we will see this threshold also plays an integral role in risk-averse case.

Theorem 2.1.5. The (conditional) expected excess payoff of investing in the risky asset is
$$E(\alpha T - L|N_t)$$
where $\alpha = r - r_0$ is the excess return.

- The expected excess payoff is negative if $t < t^*$, and the investor will only invest in the risk-free asset.
- The expected excess payoff is positive if $t > t^*$, and the investor will fully invest in the risky asset.
2.2 Risk-averse case

In this section, we extend the model to the case of risk-averse agents. The first important result it generates is that agents increase their holding gradually instead of bang-bang adjustment (from zero holding of the risky asset to full investment). This section will consider agents with a strictly increasing and concave utility function, \( U(C) \), characterize the optimal allocation, and prove that if and only if \( t > t^* \), the agents will hold a positive amount of the risky asset. We will study how changes in the risk-free rate (e.g. through monetary policy) would affect risk-taking of investors, that is we will take the comparative statics with respect to the risk-free rate (holding risk premium \( \alpha_0 \) constant).

In the risk-neutral case, only the expectation is of concern. However, in the risk-averse case, not only the expectation, but the whole distribution is of concern. Thus this section begins by proving the first-order stochastic dominance of \( T|\mathcal{N}_t \), which is a stronger version of Theorem 2.1.1.

**Theorem 2.2.1** (First-order stochastic dominance on waiting time).

\[ t_1 < t_2 \Leftrightarrow T^{-1}|\mathcal{N}_{t_1} \succ T^{-1}|\mathcal{N}_{t_2} \Leftrightarrow T|\mathcal{N}_{t_1} \succ T|\mathcal{N}_{t_2} \]

**Proof.** Let \( G_t(y) \) denote the CDF of \( T^{-1}|\mathcal{N}_t \).

Once again, we can check the conditions for switching the order of integration and differentiation are satisfied, then

\[
\frac{\partial G_t(y)}{\partial t} = \frac{\partial}{\partial t} \int \lambda \pi(\lambda)e^{-\lambda t}d\lambda \int_0^\infty \lambda e^{-\lambda x}dx d\lambda
\]

\[
= \frac{\partial}{\partial t} \int \lambda \pi(\lambda)e^{-\lambda t} \frac{1}{\lambda} e^{-\lambda y} d\lambda
\]

\[
= \frac{\partial}{\partial t} \int \lambda \pi(\lambda)e^{-\lambda t} e^{-\frac{\lambda y}{2}} d\lambda
\]

\[
= \int \lambda \pi(\lambda)e^{-\lambda t} (\pi(\lambda)e^{-\lambda t} e^{-\lambda t} d\lambda) + \pi(\lambda)e^{-\lambda t} \pi(\lambda)e^{-\lambda t} \lambda d\lambda
\]

\[
= -E_t(\lambda \cdot e^{-\frac{\lambda y}{2}}) + E_t(\lambda) \cdot E_t(e^{-\frac{\lambda y}{2}})
\]

\[
= -\text{cov}(\lambda, e^{-\frac{\lambda y}{2}}|\mathcal{N}_t) \geq 0.
\]

The second equivalence can be proved similarly or via property of first order stochastic dominance.

\[ \square \]
The above theorem means that as time passes by without the asset blowing up, our posterior distribution of $T$ shifts towards the right (as a whole as opposed to just the mean shifting towards the right). Thus even for risk averse agent, the asset becomes more attractive.

We now characterize the optimal allocation. At any time $t$, the agent maximizes the following objective function:

$$
\max_p \int_\tau U \left[ r_0 + p \cdot \left( \alpha - \frac{L}{\tau} \right) \right] \frac{\tau f_i(\tau)}{E_i(T)} d\tau.
$$

(2.5)

Note that the $E_i(\tau)$ is given at time $t$ and is not a function of $p$ or $\tau$, and can be taken out of the integral. Hence it is equivalent to maximize the following objective function:

$$
\max_{p_t} \int_\tau U \left[ r_0 + p_t \cdot \left( \alpha - \frac{L}{\tau} \right) \right] \tau f_i(\tau)d\tau.
$$

(2.6)

To sum up, the optimal solution is characterized by

- At interior solution, the FOC characterizing optimal allocation is
  $$
  \int_\tau U'' \left[ r_0 + p_t \cdot \left( \alpha - \frac{L}{\tau} \right) \right] \left( \alpha - \frac{L}{\tau} \right) \tau f_i(\tau)d\tau = 0.
  $$
  (2.7)

- If $p_t^* = 0$, then
  $$
  \int_\tau U'' \left[ r_0 + p_t \cdot \left( \alpha - \frac{L}{\tau} \right) \right] \left( \alpha - \frac{L}{\tau} \right) \tau f_i(\tau)d\tau \leq 0.
  $$
  (2.8)

- If $p_t^* = 1$, then
  $$
  \int_\tau U'' \left[ r_0 + p_t \cdot \left( \alpha - \frac{L}{\tau} \right) \right] \left( \alpha - \frac{L}{\tau} \right) \tau f_i(\tau)d\tau \geq 0.
  $$
  (2.9)

Recall that $t^*$ is the indifference point in the risk-neutral case–if more than $t^*$ passes without a disaster, the agent would be better off switching to the unfamiliar asset. This is also the point separating the corner solution from the interior solution.

**Theorem 2.2.2.** When $t < t^*$ it is optimal to have $p = 0$; when $t > t^*$ it is optimal to have $p > 0$. 
Proof. We look at the derivative with respect to \( p_t \) evaluated at \( p_t = 0 \):

\[
\int_{\tau} U' \left[ r_0 + p_t \cdot \left( \alpha - \frac{L}{\tau} \right) \right] \left( \alpha - \frac{L}{\tau} \right) \tau f_1(\tau) d\tau \bigg|_{p_t=0} = U'(r_0) \int_{\tau} (\alpha \tau - L) f_1(\tau) d\tau
\]

Obviously, when and only when \( t > t^* \), the expression is positive.

I now explore how a decrease in interest rate will affect risk-taking.

**Theorem 2.2.3.** With a quadratic utility function, at any time \( t \), a reduction in the interest rate will always increase risk-taking. With a positive third derivative \( U'''(\cdot) > 0 \), the result is ambiguous.

**Proof.** We first consider a corner solution: We know the threshold is determined by the relative size of \( E(T|N_t) \) and \( \frac{L}{r-r_0} \). Since \( E(T|N_t) \) is increasing in \( t \) and does not change with \( r_0 \), and \( \frac{L}{r-r_0} \) increases with \( r_0 \). A decrease in \( r_0 \) will result in a lower threshold \( t^* \) for every prior. Hence, assuming a continuum of heterogeneous priors, a lower interest rate will induce some agents to buy the asset.

We then consider the comparative statics when we have an interior solution:

We can take the total derivative of (2.7) with respect to \( r_0 \) and set it to zero. Solving for \( p'_t(r_0) \), we have:

\[
p'_t(r_0) = -\frac{\int_{\tau} U'' \left[ r_0 + p_t \cdot \left( \alpha - \frac{L}{\tau} \right) \right] (\alpha \tau - L) f_1(\tau) d\tau}{\int_{\tau} U'' \left[ r_0 + p_t \cdot \left( \alpha - \frac{L}{\tau} \right) \right] \left( \alpha - \frac{L}{\tau} \right)^2 \tau f_1(\tau) d\tau}.
\] (2.10)

The denominator is clearly negative. The numerator is more complicated.
where $\Delta U''(\cdot) = U''(\cdot) - U''(r_0)$. With a quadratic utility function, the second and third terms go to zero and hence the whole expression is negative.

With positive $U'''(\cdot)$, however, the last two terms are positive and the sign of the whole expression is ambiguous.

The intuition for the result in Equation (2.13) is as follows. As $r_0$ increases, the whole wealth level increases for every state and the marginal utility decreases. Marginal utility (of consumption) drops in every state: in good states $\alpha > \frac{L}{\tau}$, marginal utility of increasing $p$ drops; in bad states, marginal disutility of increasing $p$ decreases as the reduced marginal utility of consumption makes the monetary loss less bad from a utility perspective. Both of the effects are weighted by the excess payoff in each state (disutility simplify are weighted by negative weights). Assuming the marginal utility (of consumption) drops by the same amount ($U''(r_0)$) for every state (we will correct for positive $U'''(\cdot)$ later), the first effect (from good states) dominates because the expected excess payoff is positive. Thus the overall marginal utility (from increasing $p$) drops below zero, increasing $p$ decreases the expected payoff, and this makes the $p$ to go down as $r_0$ increases.

The third derivative’s effects go the other way. In bad states, marginal utility of consumption decreases more than $U''(r_0)\Delta U'' [r_0 + p \cdot (\alpha - \frac{L}{\tau})] < 0$ and the monetary loss in those states is attenuated even more (strengthening the second effect). In good states, marginal utility of consumption does not decrease as much, and hence the utility gain from monetary gain is even greater than the second order approximation. In general, a positive $U'''(\cdot)$ makes the downside risk less frightening and the upside gain even more attractive (compared to second order approximation).

The relative magnitudes of these two forces and hence the overall effect of $r_0$ depend on the utility function. Table 2.1 summarizes the decomposition of effects of an increase in risk free rate.

**Example** ($q'_t(r_0) < 0$). A quadratic utility function $U(C) = a_1(C - a_2)^2 + a_3$ where $a_1 < 0$ and $a_2 > r_0 + \alpha$ will imply $q'_t(r_0) < 0$.

**Example** ($q'_t(r_0) = 0$). A CARA utility function $U(C) = -\exp(-a_1 C)$ where $a_1 > 0$ will imply $q'_t(r_0) = 0$. 


Chapter 2: The Portfolio Allocation Decision

Good States | Bad States
---|---
Effect of $U'' < 0$ | gain is less attractive decrease $p$ | loss is bad increase $p$
Effect of third derivative | $U''$ is higher than quadratic case attenuate the effect of $U'' < 0$ | $U''$ is lower than quadratic case strengthen the effect of $U'' < 0$

Table 2.1: Summary of the Effect of a Reduction of Risk Free Rate

Example ($q_t'(r_0) > 0$). A utility function in the following form will yield $q_t'(r_0) > 0$

$$U(C) = \begin{cases} a_1(C - a_2)^2 + a_3 & C \leq a_2 \\ a_3 & \text{otherwise} \end{cases}$$

where $a_1 < 0$, and $a_2 \in (r_0, r_0 + \alpha)$.

Figure 2.2 summarizes the dynamics of the investor demand. The blue solid line depicts the bang-bang adjustment of a risk neutral agent. The red dashed line depicts the more smooth adjustment of a risk-averse agent. Finally, the orange dotted line depicts the effect of a reduction in risk-free rate, and compared to the red dashed line, the investor not only started investing in the unfamiliar asset earlier, but also (under certain conditions such as quadratic utility function) consistently invests higher amount.

Figure 2.2: Summary of the Dynamics of Investor Demand
2.3 Connection to behavioral finance

One motivation behind this model is to model “under-reaction and over-reaction” formally without losing tractability. In this Bayesian setting, we could incorporate such properties via the updating function, that is for the posterior, instead of using (2.1), we can use the more general form:

\[
\pi(\lambda|N_t) = \frac{\pi(\lambda)e^{-\lambda s(t)}}{\int \pi(\lambda)e^{-\lambda s(t)}d\lambda}
\]

(2.14)

where \(s(t)\) is the updating function. If \(s(t) = t\) then it is our usual “rational” Bayesian updating. We could also let \(s(t)\) look like the function is Figure 2.3, where \(s_0\) is the threshold for “under-reaction” and “over-reaction”. With suitable data and some further assumptions, we can test the hypothesis that the implied updating function is the identity function (that is the investor is rationally updating in a Bayesian fashion). Section 3.2 discusses the empirical implications of the updating function.
Chapter 3

General Equilibrium and Empirical Implications

3.1 Incorporating price

The previous section has considered the partial equilibrium where the default premium is taken to be exogenous. In this section, we consider the problem in a general equilibrium setting, where the default premium will be endogenous to clear the market where the risky asset is in fixed supply. We consider a market of heterogeneous investors, faced with these two assets, and deciding whether to hold the risky assets. We endogenize the default premium of the risky asset to clear the market. We investigate the trajectory of the default premium in this setting. In the process, we will prove two “no-trade” theorem in two simple special cases.

For simplicity, we assume that the agent to be risk-neutral, has unit mass, and that each investor has one dollar to invest between two infinitely divisible assets. We assume each investor has a prior on $\lambda$ in the form of $\pi_i(\lambda) \sim \frac{1}{n_0} Gamma(n_{0,i}, \mu_{0,i})$, but each investor has different hyper-parameters $n_{0,i}$ and $\mu_{0,i}$. $\mu_{0,i}$ specifies the agents’ prior expectation of $\lambda$, while $n_{0,i}$ specifies how confident the agent is in his prior belief.

On the supply side, we first assume the supply of the risky assets are perfectly inelastic (we will later relax this assumption), and there are $p^* < 1$ units supplied in total while the supply of the risk-free assets is perfectly elastic (it is simply the residue amount).
We will assume the following cash-flow: both assets cost one dollar per unit to purchase. When the risky asset blows up, it does not return the one dollar it costs, thus it is equivalent to losing one dollar (it is just losing the one dollar up-front). Hence we have essentially fixed $L = 1$. It is the $\alpha$ that serves the function of the price.

This simplification does not sacrifice any generality. Both $L$ and $\alpha$ are relevant for the investment decision. However, in the risk-neutral case, it is easy to see that it is redundant to have two parameters: doubling $L$ and $\alpha$ will always yield the same decision, to buy risky asset, to buy safe asset or to be indifferent. To see this, note that all that is relevant for the decision is the sign of $E(\alpha T - L)$, since we will assume $\alpha$ to be always positive, all that matters is the ratio $\frac{L}{\alpha}$.

In each time period, different investor will have a threshold $\alpha_{t,i}$ at which he is indifferent between purchasing the risky asset or the risk-free one, and there will be a $\alpha^*_t$ that will clear the market at time $t$.

**Theorem 3.1.1.** Given any continuous distribution of hyper-parameters, there exists a $\alpha^*_t$ that will clear the market at any time $t$.

**Proof.** As we show in Lemma 3.1.2, the individual threshold is $\alpha_{t,i} = \frac{n_0 \mu_{0,i} - 1}{n_0 + t}$ and since $n_{0,i}$ and $\mu_{0,i}$ has continuous distribution, so does $\alpha_{t,i}$, hence the CDF of $\alpha_{t,i}$ is continuous. We can apply Intermediate Value Theorem to obtain existence of the equilibrium price.

**Remark.** Indeed we can prove a weaker result for any continuous distribution of priors, not just for a single class of priors. In this setting, we can only prove that the market will clear at all but finitely many $t$’s. The technicalities yield no additional insights, however, so I will not prove the more general result.

**Lemma 3.1.2.** For an investor with the following prior on $\lambda$, $\pi(\lambda) \sim \frac{1}{n_0} \text{Gamma}(n_0 \mu_0)$, with $n_0 \mu_0 > 1$ the threshold $\alpha = \frac{n_0 \mu_0 - 1}{n_0 + t}$.

**Proof.** We show in Section A.1 that given such a prior, if the asset has not blown up at time $t$, then the investors’ posterior belief of $T$ follows $\frac{n_0 + t}{n_0 \mu_0} \text{F}(2, 2n_0 \mu_0)$. Then the posterior

---

1 Even in the risk-averse case, the two parameters are redundant, and scaling up or scaling down both parameters is equivalent to changing units of account.

2 This simply means that the measure of any given prior is zero.
expectation of $T$ is $\frac{n_0 + t}{n_0 \mu_0 - 1}$.
The investor is indifferent if and only if the following holds:

$$\frac{n_0 + t}{n_0 \mu_0 - 1} = L$$

When all the investors share one of the hyper-parameters $n_0$ or $\mu_0$, the equilibrium price becomes very easy to compute, all that matters is the $q^*$ quantile of the other hyperparameter. Furthermore, there is “no trade” as time passes by—the investors who started holding these risky assets will continue to hold those risky assets.

**Theorem 3.1.3.** We assume that each investor has a prior on $\lambda$ in the form of $\pi(\lambda) \sim \frac{1}{n_0} \cdot \text{Gamma}(n_0, \mu_0)$, and that each investor has the same $n_0$ but each investor has different $\mu_{0,i}$. Let $\mu_0^*$ be the $q^*$th quantile of the $n_0$’s, where $p^*$ is the quantity of risky assets supplied. Then we will have

$$\alpha_i^* = \frac{n_0 \mu_0^* - 1}{n_0 + t}$$

*Proof.* This follows from Lemma 3.1.2.

**Theorem 3.1.4.** We assume that each investor has a prior on $\lambda$ in the form of $\pi(\lambda) \sim \frac{1}{n_0} \cdot \text{Gamma}(n_0, \mu_0)$, and that each investor has the same $\mu_0$ but each investor has different $n_0$. Let $n^*$ be the $q^*$th quantile of the $n_0$’s, where $p^*$ is the quantity of risky assets supplied. Then we will have

$$\alpha_i^* = \frac{n^* \mu_0 - 1}{n^* + t}$$

*Proof.* Assume $n_{0,1} > n_{0,2}$, then we have

$$\frac{n_{0,1} \mu_0 - 1}{n_{0,1} + t} - \frac{n_{0,2} \mu_0 - 1}{n_{0,2} + t} = \frac{(1 + t \mu_0)(n_{0,1} - n_{0,2})}{(n_{0,1} + t)(n_{0,2} + t)} > 0, \forall t$$

That is the threshold default premium of investor one is always higher than that of investor two.

*Remark.* While less obvious than the previous theorem, this theorem is also quite intuitive. Since any two agents observe the same events, an agent starting with a noisier prior on $\lambda$ will continue to have a noisier posterior of $\lambda$. Ceteris paribus (meaning the same prior mean), the posterior mean on stopping time $T$ will thus always be shorter.
Chapter 3: General Equilibrium and Empirical Implications

In fact, the no-trade theorem could be generalized into a probabilistic “no trade” theorem.

**Lemma 3.1.5.** We assume that the agent to be risk-neutral, has unit mass, and that each investor has one dollar to invest. We assume each investor has a prior on $\lambda$ in the form of $\pi_i(\lambda) \sim \frac{1}{n_0} \text{Gamma}(n_{0,i}\mu_{0,i})$, but each investor has different hyper-parameters $n_{0,i}$ and $\mu_{0,i}$. Then for any two investor with hyperparameter pairs $(n_{0,1}, \mu_{0,1}), (n_{0,2}, \mu_{0,2})$, we have the following: there exists $\bar{T}(n_{0,1}, \mu_{0,1}, n_{0,2}, \mu_{0,2})$

\[ (\alpha_{t,1} - \alpha_{t,2})(\alpha_{t,2,1} - \alpha_{t,2,2}) > 0, \forall t_1, t_2 > \bar{T} \tag{3.1} \]
\[ (\alpha_{t,1} - \alpha_{t,2})(\alpha_{t,1,1} - \alpha_{t,1,2}) > 0, \forall t_1, t_2 < \bar{T} \tag{3.2} \]
\[ (\alpha_{t,1} - \alpha_{t,2})(\alpha_{t,1,1} - \alpha_{t,1,2}) < 0, \forall t_1 < \bar{T} < t_2 \tag{3.3} \]

that is, from $\bar{T}$ on the relative size of $\alpha_{t,1}$ and $\alpha_{t,2}$ remains the same.

**Proof.**

\[
\frac{n_{0,1}\mu_{0,1} - 1}{n_{0,1} + t} - \frac{n_{0,2}\mu_{0,2} - 1}{n_{0,2} + t} = \frac{(n_{0,1}\mu_{0,1} - n_{0,2}\mu_{0,2})t + (n_{0,1} - n_{0,2}) + n_{0,1}n_{0,2}(\mu_{0,1} - \mu_{0,2})}{(n_{0,1} + t)(n_{0,2} + t)}
\]

The denominator is clearly positive. If $n_{0,1}\mu_{0,1} - n_{0,2}\mu_{0,2} \neq 0$, then choose

\[ \bar{T} = \| \frac{(n_{0,1} - n_{0,2}) + n_{0,1}n_{0,2}(\mu_{0,1} - \mu_{0,2})}{n_{0,1}\mu_{0,1} - n_{0,2}\mu_{0,2}} \|, \]

then the first term of the numerator always dominate. Otherwise, the first term disappears and choose $\bar{T} = 0$. \qed

The key implication of the previous lemma is that the relative size of $\alpha_{t,i}$ between any two investors switch only once! Consequently we have a probabilistic “no trade” theorem:

**Theorem 3.1.6.** We assume that there are finitely many risk-neutral agents, and that each investor has one dollar to invest.\(^3\) We assume each investor has a prior on $\lambda$ in the form of $\pi_i(\lambda) \sim \frac{1}{n_0} \text{Gamma}(n_{0,i}\mu_{0,i})$, but each investor has different hyper-parameters $n_{0,i}$ and $\mu_{0,i}$. Then we have: Let $I_t$ be the indicator whether there is trade at time $t$. For any $t_0$ and $\Delta t$,

\[ \int_{t_0}^{t_0 + \Delta t} I_t \, dt = 0 \]

\(^3\)We can easily extend this to countably infinite case, but that seems a useless extension.
Proof. From Lemma 3.1.5, we have a very loose upper bound on the number of trades: \( \binom{n}{2} \), where \( n \) is the number of investors.

We now give some numerical examples:

**Example.** Assume there is a unit mass of investors, each with the same amount to invest. Normalize the total amount of money that could be invested to 1. Suppose there is 0.5 unit of risky assets perfectly inelastically supplied. We will generate the price path of \( \alpha_t \) such that the market will always clear.

1. Suppose for each investor’s prior, \( n_0 = 5 \) (point mass distribution), and \( \mu_0 \sim LN(0, 1) \), then \( \alpha_t = \frac{4}{5+t} \).

2. Suppose for each investor’s prior, \( n_0 \) has a point mass, and \( \mu_0 \) is distributed in such a way that its median is \( \hat{\mu}_0 \). then \( \alpha_t = \frac{n_0 \hat{\mu}_0 - 1}{n_0 + t} \).

3. Finally we simulate the following when \( n_0 \) comes from a discrete uniform over \( \{1, 2, 3, 4, 5\} \), and \( \mu_0 \) is independently distributed as uniform over \([1, 3]\). Figure 3.1 presents the simulated premium trajectory.

**Theorem 3.1.7.** The market clearing default premium monotonically decreases: \( \alpha_t^* \leq \alpha_t^{*'} \iff t_1 \geq t_2 \).

**Proof.** Let \( I_i(\alpha_t, t) \) be the indicator function of \( E_i(T|N_t = 0) \geq \frac{I}{\alpha_t} \), that is the investor is willing to purchase the risky asset given the default premium. We know that the indicator function is monotonically increasing in \( \alpha_t \) and \( t \). We know at market clearing \( \alpha_t^* \), we have

\[
d(\alpha_t, t) = p^*
\]

where the demand function at time \( t \) is defined as:

\[
d(\alpha_t, t) := \int I_i(\alpha_t, t) d\mu(i)
\]

We know immediately that \( d(\alpha_t, t) \) is also monotonically increasing in \( \alpha_t \) and \( t \). Suppose we have \( \alpha_t^* < \alpha_t^{*'} \) with \( t_1 < t_2 \), we will have

\[
d(\alpha_t^*, t_1) < d(\alpha_t^*, t_2) < d(\alpha_t^{*'}, t_2)
\]

Contradiction! Therefore, \( \alpha_t^* \leq \alpha_t^{*'} \iff t_1 \geq t_2 \).
Figure 3.1: Trajectory of the default premium
We can now generalize the supply of the unfamiliar asset. Let \( s(\alpha) \) be the supply function of the unfamiliar asset. It is monotonically decreasing in \( \alpha \) with the following constraints:

\[
\lim_{\alpha \to 0} s(\alpha) = \infty \quad (3.5)
\]
\[
\lim_{\alpha \to \infty} s(\alpha) = 0 \quad (3.6)
\]

that is there is more of this kind of assets created if the default premium is low, and if the default premium disappears, there will be infinite supply of this asset as it provides a perfect arbitrage opportunity.

We will still have the existence of market clearing default premium:

**Theorem 3.1.8.** Given any continuous distribution of hyper-parameters, a monotonically decreasing supply function \( s(\alpha) \) that satisfy Equations 3.53.6, there exists a \( \alpha^*_t \) that will clear the market at any time \( t \).

**Proof.** We look at excess demand function \( ed(\alpha_t) = d(\alpha_t) - s(\alpha_t) \) and apply the intermediate value theorem. □

Theorem 3.1.4 and Lemma 3.1.2, 3.1.5 still hold as their proof does not rely on the specification of the supply function.

We can also study the dynamics of default premium and quantity supplied of the unfamiliar asset.

**Theorem 3.1.9.** As time passes by, the market clearing \( \alpha^*_t \) monotonically decreases and the quantity supplied/demanded monotonically increases.

**Proof.** For any \( t_1 < t_2 \) we know from Theorem 3.1.4 that \( d(\alpha, t_1) < d(\alpha, t_2) \) for any \( \alpha \). Suppose there exists \( t_1 < t_2 \) such that \( \alpha^*_1 < \alpha^*_2 \), then we have

\[
d(\alpha^*_1, t_2) > d(\alpha^*_1, t_1) = s(\alpha^*_1) > s(\alpha^*_2) = d(\alpha^*_2, t_2)
\]

Contradicting the fact that \( d(\alpha, t) \) is increasing in \( \alpha \). Thus we conclude the \( \alpha_t \) monotonically decreases. By the monotonicity of the supply function, we conclude that the quantity supplied monotonically increases. □
In essence this theorem says that as the investor view the unfamiliar assets to be safer, their demand increases, and this drives up the amount supplied as well. Thus we will expect to see not only an increase in price (a decrease in default premium), but also an increase in the amount of the assets. This is qualitatively consistent with the observation in the recent financial crisis that, not only has the price of CDO decreased, their supply exploded.

3.2 Empirical Implications

In this section, we discuss what the theories would imply about the real world data we observe. The ideal data to look at will be data on fixed-income assets such as below investment grade corporate bonds, arguably an example of “unfamiliar assets”. The cash flow structure of such bonds are similar to the risky asset described in the model. Similarly, we could expect CDO data to exhibit similar properties. http://www.library.hbs.edu/go/mergentfisd.html contains data we will be looking at. Unfortunately, we are unable to obtain this dataset due to the proprietary nature of such data, and thus the discussed methods will not be carried out.

One approach is to look at the difference of the inverse of the default premiums of the same bonds at different times, amenable to a simple reduced-form linear regression model. This approach is consistent if all of the following conditions hold:

- The only relevant information is historical performance.
- There is no trade involving the marginal buyer.
- The marginal buyer has the following prior on $\lambda$, $\pi(\lambda) \sim \frac{1}{n_0} Gamma(n_0\mu_0)$, with $n_0\mu_0 > 1$.
- The marginal buyer updates his prior in a perfectly rational Bayesian fashion.

We know from Lemma 3.1.2 that the marginal buyer’s threshold $\alpha_t = \frac{n_0\mu_0 - 1}{n_0 + t}$. Since he is always the marginal investor, then his threshold will be the market clearing premium. We have the surprisingly simple relation: the inverse of the premiums of the same asset $\alpha_t$. 


and $\alpha_t$ at two given times $t_1$ and $t_2$ will follow the following exact relation:

$$\alpha_{t_2}^{-1} = \alpha_{t_1}^{-1} + \frac{t_2 - t_1}{n_0 \mu_0 - 1}$$  \hspace{1cm} (3.7)

This relation holds regardless of the hyper-parameters, because the premium $\alpha_t$ already incorporates this information. In equation (3.7), the difference of the inverse of default premiums of the same stock at different times (interpreted as price) is linear the length of time between two price measures.

Motivated by the discussion above, we can regress the price (the inverse of default premium) at time $t_2$ on the price at time $t_1$, the difference in $t_2$ and $t_1$:

$$p_{i,2} = \beta_0 + \beta_1 p_{i,1} + \beta_2 (t_{i,2} - t_{i,1}) + \epsilon_i$$ \hspace{1cm} (3.8)

where $p_{i,t} = \alpha_{t,i}^{-1}$.

According to the theory, we would expect both $\beta_1$ and $\beta_2$ to be positive. Moreover, $\beta_1$ will be close to unity if $\epsilon_i$ is conditionally (on $x$) independent of $t_{i,2} - t_{i,1}$ and $p_{i,1}^{-1}$. This is a very tricky condition to interpret. After all, the future information flow that is relevant to the asset performance is a stochastic process, our prior on which will determine $p_{i,1}$, and we would expect our prior knowledge of the stochastic process to be not independent of the actual stochastic process. Note that Equation (3.8) can be decomposed into three parts: the first part $\beta_1 p_{i,1}^{-1}$ encodes all the information up to time $t_1$, the second part is the drift, and while the third part encodes all the information other than historical performance between time $t_1$ and $t_2$. The condition means that if all the information (other than historical performance) we gathered from $t_1$ to $t_2$ is observed at time $t_1$ instead, the marginal investor will still have the threshold premium $p_{i,1}$ in expectation.\footnote{More accurately, the expectation of the inverse of the premium will not change in expectation.} In some sense, the market is “informationally efficient”, and the price is a martingale with respect to the filtration induced by all the non-historical information.

Note that Equation 3.8 implies that the price (the inverse of the premium) follows a random walk with a (positive) drift. The drift is due to the fact that we have chosen all the assets that has survived up to time $t_{i,2}$—the impeccable track record which is the result of selection bias, leads to the lower premium at a later time. Such biases caused by truncation...
tion by survivorship is studied in the “survivorship bias” literature, such as Ball and Watts (1979); Brown et al. (1992).

An alternative approach relaxes the last condition in the previous approach. It is more also “structural” in the sense that it estimates the parameter of the updating function, introduced in Section 2.3, and we can allow for such behavioral deviations and test for the existence of such behavioral irrationality.

First we restrict the set of updating function \( s(t) \) to the following class: \( s(t) = (t_0^{t_0})^\alpha \). At \( t = t_0 \), the updating function will cross the 45 degree line for the first time since \( t > 0 \). If “under-reaction” and “over-reaction” do play a role in actual market, we would expect \( \alpha > 1 \)—Before \( t_0 \), it is the domain of under-reaction, and after \( t_0 \), it is the domain of “over-reaction”.

For actual estimation, we first take the difference of the premiums at the those two times.

\[
p_{i,2}^{-1} - p_{i,1}^{-1} = \frac{t_{i,2}^\alpha - t_{i,1}^\alpha}{t_0^{\alpha - 1}(n_0\mu_0 - 1)}
\]

Hence, we have the following estimation function:

\[
p_{i,2}^{-1} = \beta_0 + \beta_1 p_{i,1}^{-1} + \beta_2(t_{i,2}^\alpha - t_{i,1}^\alpha) + \varepsilon_i
\]

The above equation is similar to Equation 3.8 except the two \( t \)'s are raised to the \( \alpha \) power. This could not be estimated via a linear regression model as it is non-linear in the parameters, but standard methods such as MLE can be used to estimate this model.

Equation (3.9) is similar to Equation (3.8) except that the deterministic drift is not linear in time difference. If \( \alpha > 1 \), as will be predicted by the “under-reaction” and “over-reaction” literature, that is the updating function is convex in \( t \), given the same time difference, the higher \( t_{i,1} \) the higher the drift term. This is because as time progresses, investors over-update and given the successful performance (by assumption of survival) they tend to revise their estimates of the blow-up frequency downward more rapidly.
Chapter 4

Conclusion

This paper started by modeling a simple learning process of an investor about an unfamiliar asset, and characterized the investor’s demand for an unfamiliar asset in this learning framework—under quite general conditions (nondogmatic prior belief), given an exogenous default premium, there will be a threshold time $t^*$ after which the investor will start investing in the asset and remain invested in the asset. The paper then endogenized the default premium to clear the market, and characterize the default premium trajectory—even with the same fundamentals, the default premium monotonically decreases to zero, and the quantity supplied/demanded monotonically increases as time passes by (assuming no crash happens). In other words, purely by chance, an asset might continue performing for an extended period of time, and because of this track record, not only have investors bid up its price, but also have purchases larger and larger quantities of this asset. When it does crash, a significant amount of wealth will be destroyed (or transfered).

One implication of the model is that learning, innocuous as it might sound, could be destabilizing. This suggests that policy makers should be more cautious when it comes to financial liberalization. Liberalizing counties often encounter some form of crisis subsequently partly because liberalization exposes existing problems, decreases the franchise value of a bank liscence, and it is often poorly sequenced. Caprio and Honohan (1999); Caprio et al. (2006); Honohan (n.d.) This paper suggests another channel. When financial liberalization takes place, the whole financial system becomes unfamiliar, inducing the investors begin to learn about all the assets in this new system, and this destabilizing force
might create another boom and bust cycle.
Appendix A

Theorems and Proofs

A.1 A Note on Computation

One of the practical difficulties with the Bayesian approach is its lack of computational tractability. In this section, we propose a specification that allows more tractability.

**Theorem A.1.1.** Given

\[
\pi(\lambda) \sim \frac{1}{n_0} \text{Gamma}(n_0\mu_0) \tag{A.1}
\]

the updating rule

\[
\pi(\lambda | N_t) = \frac{\pi(\lambda) e^{-\lambda s(t)}}{\int \pi(\lambda) e^{-\lambda s(t)} d\lambda} \tag{A.2}
\]

and

\[
T | \lambda \sim \frac{1}{\lambda} \text{Expo} \tag{A.3}
\]

Then,

\[
\lambda | F_t \sim \frac{1}{n_0 + t} \text{Gamma}(n_0\mu_0) \tag{A.4}
\]

and,

\[
T | N_t \sim \frac{n_0 + t}{n_0\mu_0} F(2, 2(n_0\mu_0)) \tag{A.5}
\]

We will expound what the theorem means in words: The prior will be in the form of

\[
\lambda \sim \frac{1}{n_0} \text{Gamma}(n_0\mu_0) \sim [\mu_0, \frac{1}{n_0}].
\]

That is we are perfectly free to adjust our prior in terms of the two parameters of our Gamma distribution with \(n_0\) corresponding to how certain we are, and \(\mu_0\) corresponding to our prior mean.
When time $t$ passes with no blow-up, we form our posterior, which is still a Gamma Distribution—$\lambda | F_t \sim \frac{1}{n_0 + t} \Gamma(n_0 \mu_0) \sim \left[ \frac{n_0}{n_0 + t} \mu_0, \frac{n_0 \mu_0}{(n_0 + t)^2} \right]$. This gives us the first layer of tractability—posterior closure. Now we consider what really matters for the agent—the blow-up time $T$. As noted, $T | \lambda \sim \frac{1}{\lambda} \text{Expo}$. The distribution of $T$ (unconditional on $\lambda$, but conditional on all information about $\lambda$, that is the posterior), is a mixture distribution. Continuing with our specification $\lambda | N_t \sim \frac{1}{n_0 + t} \Gamma(n_0 \mu_0)$, we will have $T | N_t \sim \frac{n_0 + t}{n_0 \mu_0} F(2, 2(n_0 \mu_0))$. While the $F$-distribution is much more complicated than other distributions, it is in the Natural Exponential Family, and its many properties (like raw moments, and characteristic function) are well-known. It is at least in closed-form so that no numerical method needs to be employed in this case.

We will prove the theorem in parts: We first state the conjugacy of Gamma distribution and Poisson distribution.

**Lemma A.1.2.** Given the prior on $\lambda$:

$$\pi(\lambda) \sim \frac{1}{n_0} \text{Gamma}(n_0 \mu_0) \quad (A.6)$$

and observation $x_i \overset{iid}{\sim} \text{Poisson}(\lambda)$ ($i = 1, 2, \ldots, n$), then the posterior on $\lambda$ is

$$\pi(\lambda | \{X_i\}) \sim \frac{1}{n_0 + n} \text{Gamma}(n_0 \mu_0 + n \cdot \bar{X}) \quad (A.7)$$

where $\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$.

**Proof.** This is a standard result in statistics concerning Natural Exponential Family (Canonical Exponential Family, or linear Exponential Family). See

**Corollary A.1.3.**

$$\pi(\lambda) \sim \frac{1}{n_0} \text{Gamma}(n_0 \mu_0) \quad (A.8)$$

the updating rule:

$$\pi(\lambda | N_t) = \frac{\pi(\lambda) e^{-\lambda s(t)}}{\int \pi(\lambda) e^{-\lambda s(t)} d\lambda} \quad (A.9)$$

Then,

$$\lambda | F_t \sim \frac{1}{n_0 + t} \text{Gamma}(n_0 \mu_0) \quad (A.10)$$
Proof. This is because that if there is no blow up until time $t$, it is equivalent to observing $\text{Poisson}(\mu t) = 0$. 

Now we move on to the mixture distribution of Gamma Distribution and Exponential Distribution.

**Lemma A.1.4.** Given marginal distribution on $\lambda$ is:

$$\mu \sim b \cdot \text{Gamma}(a) \quad \text{(A.11)}$$

and the conditional distribution of $T|\lambda$:

$$T|\lambda \sim \frac{1}{\lambda} \text{Expo} \quad \text{(A.12)}$$

we have the mixture distribution (unconditional distribution) of $T$ to be:

$$T \sim \frac{1}{ab} F(2, 2a) \quad \text{(A.13)}$$

Proof. The actual derivation is tedious and calculation-intensive, but it is easy to verify the claim using the characteristic function with the help of Mathematica. Since the $F$ distribution does not have Moment generating function, we have to rely on the Characteristic function. The idea is that if the characteristic functions of two distributions are the same these two distributions are the same. See the attached Mathematica verification.

Then Theorem A.1.1 is a simple consequence of the previous two lemmas.

**A.2 Proof of Lemma A.1.4**
Verifying F distribution as Gamma Mixture Distribution of Exponential Distribution

This is the Characteristic Function of $\frac{1}{ab}F(2,2a)$ distribution

In[1]:= FullSimplify[CharacteristicFunction[FRatioDistribution[2, 2 a], $\frac{1}{ab} s$]]
Out[1]= $a ã- ä s
b ExpIntegralE\[1 + a, -\frac{i s}{b}\]

The following code calculates the Gamma Mixture Distribution of Exponential Distribution.

In[2]:= mix = ParameterMixtureDistribution[ExponentialDistribution[\[Lambda]], \[Lambda] \[Distributed] GammaDistribution[a, b]]; The following code calculates the characteristic function of the mixture distribution:

In[3]:= CharacteristicFunction[mix, s]
Out[3]= $a ã- ä s
b ExpIntegralE\[1 + a, -\frac{i s}{b}\]
References


