Contents

Implications of Model Risk for Market Microstructure: Pricing and Trading of Illiquid Securities
Yu Wang & Donald J. Smith 704 - 720

Managing Credit Risk - a Method for Maintaining an Optimal Distance to Default
Jana Simon 721 - 730

Superannuation Fund Communications in Good and Bad Times
Malcolm Smith, Paul Gerrans, & Jacqui Whale 731 - 755

Two Cultures Shaping the Corporate Governance Debate
Donald G. Margotta 756 - 769

Revisiting Issues Involving Risk
David Eagle, Arsen Djatej, Brian Grinder, & Dean Kiefer 770 - 786

Representativeness Heuristic Can Cause Asset Price Underreaction to New Information in an Asset Market with Strategic Interactions
Guo Ying Luo 787 - 798

Data Visualization: An Alternative and Complementary Learning Strategy to Teaching Ratio Analysis
Dr. David DeBoskey & Martha Doran 799 - 821

The Credit Quality Profile of The S&P 500 Index- A Historical Analysis
Dr. Vinod Venkiteshwaran 822 - 833

The Impact of Mortgage Securitization on the Housing Bubble and Subprime Mortgage Crisis: A Self-organization Perspective
Hong Y. Park, Hyejung Chang, & Kaustav Misra 834 - 855

Commercial Banks in Investment Banking Underwriting: Certification Effect or Conflicts of Interest
Yong-Chern Su, Yu-shin Lin, & Han-Ching Huang 856 - 868

The Determinants of Capital Structure in A Tax-Free Environment: An Extreme Bounds Analysis
Imad Moosa & Wafaa Sbeti 869 – 888

www.irjaf.com
Implications of Model Risk for Market Microstructure: Pricing and Trading of Illiquid Securities\textsuperscript{1}

Yu Wang  
Financial Engineering Group  
Models, Mission and Research Division  
Freddie Mac, 1551 Park Run Drive, McLean, VA 22102  
yu_wang@freddiemac.com

Donald J. Smith  
Department of Finance and Economics  
School of Management  
595 Commonwealth Avenue  
Boston, MA 02215  
donsmith@bu.edu

\textsuperscript{1}The model and analytical results in this paper are based on a June 2007 manuscript by Yu Wang entitled ``Model Risk, Performance Evaluation, Risk Premium, and Implications on Pricing and Trading of Securities.” That paper was motivated by discussions with Joseph Langsam of Morgan Stanley.

*corresponding author
Abstract
This paper examines some of the implications of model risk for market microstructure, in particular for a financial institution making a market in illiquid securities. Using a model to measure the extent of model risk, the paper derives the bid and ask prices, the bid-ask spread, and the risk premia for bearing model risk. For the simplifying case of only two illiquid securities, numerical examples are used to illustrate that setting the bid and ask prices to compensate for model risk, given the impact on the overall balance sheet resulting from trades at those prices, can lead to intentional over-pricing and under-pricing vis-a-vis the fair value estimated by the model.
JEL Classification: G12, G14, G21.
Keywords: model risk, market microstructure, bid-ask spread.

I. Introduction
Widespread use of structured financial products and over-the-counter derivatives magnifies the importance of model risk for financial market participants. Financial institutions that act as market makers in illiquid securities and derivatives inevitably must rely on valuation models to set bid and ask prices as well as to formulate trading and hedging strategies. Financial assets and liabilities that cannot be reliably “marked to market” because of thin trading or non-public price data need to be “marked to model” in some manner. There is no doubt that model risk, which essentially is the risk that values produced by statistical models can differ significantly from their “true values”, impacts profitability.
In the U.S., the new accounting standard FAS 157 on fair value measurements requires disclosure of the manner in which valuations are obtained. Assets and liabilities that are reported at fair value on the balance sheet will be categorized according to the type of input: Level 1 inputs are directly observed prices in active markets (e.g., exchange-traded equities and derivative contracts); Level 2 inputs are observations in related markets (e.g., using current LIBOR swap rates to value a pre-existing interest rate swap); and Level 3 inputs are the institution's own assumptions to its models. Obviously, valuations lumped into the third category will have significant model risk.
Model risk can arise in various ways. Derman (1996) identifies several sources: (i) inapplicability of modeling; (ii) incorrect model; (iii) correct model, but an incorrect solution; (iv) correct model, but an inappropriate use; (v) a badly approximated solution; (vi) software and hardware bugs; and (vii) unstable data. Estimating values for thinly traded securities is hard enough in normal times when financial markets are smoothly functioning. Model risk becomes a critical factor during a liquidity crisis when a financial institution has to value securities like CDOs (collateralized debt obligations) backed by sub-prime mortgages.
Our objective in this paper is to demonstrate that model risk can be an important factor in determining bid-ask spreads and, therefore, belongs in the discussion of market microstructure.
We examine some of the implications of model risk for the pricing and trading of securities, in particular illiquid securities for which market prices are not readily available. We take the perspective of a financial institution (for instance, a commercial or investment bank) that is a two-way market maker, quoting both bid and ask prices. These securities do not trade on an exchange. Instead, transactions only take place in the over-the-counter (OTC) market. Therefore, prices at which trades occur with other market makers may not be known to some institutions or they may only be learned with a time lag. That can make it very difficult for the estimation errors on prices to be detected and corrected.
We will allow the bank in our model to take either long or short positions in the securities. As a market maker, the bank will want to set its bid and ask prices correctly to earn a risk-adjusted profit on each transaction. But the bank will also act as a proprietary trader and have a view on the market. When expecting positive returns on the securities, it will want to have more trading inventory (i.e., more long positions). When expecting negative returns, it will want to have less inventory (and maybe even short positions) along with the correct bid-ask spread. Financial institutions such as this bank rely heavily on models to trade in these markets, exposing them to significant model risk.

We start by assuming a standard mean-variance optimization model for the financial institution and introduce a metric for the extent of model risk. Then we derive the range for bid and ask prices for a transaction of a particular size in one of the securities and calculate the required risk premia for bearing model risk. To keep the model tractable, we assume that there are only two illiquid securities and that their characteristics and the bank's holdings in each are the same. Then we are able to generate several numerical examples to illustrate the implications of model risk. While we do not work with a particular security type in this paper, it might be useful to keep in mind a CDO tranche that has cash flows tied to the default and prepayment experience on a pool of residential mortgages. These typically are illiquid securities traded only in OTC markets and do require a sophisticated model for valuation. Under FAS 157, Level 3 inputs are used.

The analysis shows that the bid and ask prices, bid-ask spreads, and model risk premia depend on the values of four sets of variables: (1) the institution's current holdings of securities (i.e., long and short positions of various quantities), (2) the mean and variance of individual security returns and the correlation between the returns, (3) the mean and variance of estimation errors (arising due to model risk) and the correlation between the errors, and (4) the sizes of trades. In this model we observe over-pricing, under-pricing, or fair pricing of the illiquid securities. Over-pricing here means that the bid and ask prices are set above the valuation given by the model. Under-pricing means the bid and ask prices are set below the level indicated by the model and fair pricing means they are set approximately at the same value. Notice that if the securities were highly liquid, competitive pressures probably would not allow an individual market maker to manipulate its bid and ask prices in this manner.

When expected future values for the illiquid securities are high, over-pricing occurs in the model to produce larger trading profits. Both the bid and ask prices are raised, relative to the values produced by the model, in order to increase inventory. The bank is willing to buy more securities at a higher price and will only be willing to sell at a higher price. Likewise, low expected future values lead to under-pricing. The market maker wants to reduce its holdings, so both bid and ask prices are lowered relative to the estimated values given by the models. Model risk in this model adds additional uncertainty to the future values of assets---that makes the institution behave as if it is more risk averse. Therefore, other things being equal, long positions on the securities are more likely to lead to under-pricing to reduce holdings, whereas short positions lead to over-pricing to rebuild inventory.

These results demonstrate that model risk is directly related to the literature on market microstructure, in particular, work by Stoll (1978), Ho and Stoll (1981), and O'Hara and Oldfield (1986). These papers study how the current inventory of securities, and other factors such as the moments of returns on securities and expected order flows, would affect how a risk-averse market maker sets its bid and ask prices. Our paper differs from these in that we focus on illiquid OTC securities rather than exchanged-traded equities and we consider market risk and model risk rather than just market risk. Nevertheless, there are some clear parallels in the results. See

There is also a growing literature on model risk itself. Derman (1996) and Dowd (2006) discuss the sources of model risk and ways that it can be managed in practice. Cont (2006), Green and Figlewski (1999), Hull and Suo (2002), and Jacquier and Jarrow (2000) assess and quantify model risk for individual models that are used to price contingent claims.

II. A Model of Model Risk

This section develops a model of model risk. To keep the analysis as simple as possible, we consider a market of only two illiquid securities. First, we state all the formal assumptions that are made in building the model. Then we introduce a metric for the extent of model risk. After that, we derive the bid and ask prices for the securities and the model risk premia. Last, we discuss the effects of model risk on the pricing and trading of the securities.

II.1 Assumptions

Denote by \( \epsilon_i = c_i(x_{it}, \theta_i), i = 1, 2 \) the estimated fair values of the securities given by models at time \( t \), where \( x_i \) is a vector of state variables that the price of the \( i \)th security depends on, and \( \theta_i \in \Theta \) is a vector of parameters in the models for the \( i \)th security. For example, if the security is a mortgage-backed CDO, state variables could include measures for the real economy and credit market conditions, while the parameters could include a specific prepayment rate for the mortgage pool and the term structure of Treasury yields. Denote by \( \epsilon_i = \epsilon(x_{it}, \theta_i), i = 1, 2 \) the true fair values of the securities at time \( t \). Assume that the bank holds \( q_i \) units of the \( i \)th security at \( t \), where \( q_i \) can be either positive, or negative, or zero. Positive values for \( q_i \) indicate long positions and negative values indicate short positions. While one could interpret a short position to mean that the bank has issued the security, we mean that market maker has sold more than it purchased. It has borrowed the securities to sell them and the liability represents its obligation to return them. Alternatively, the long or short position can be in an illiquid forward market on a security that itself is liquid. The key point is that the positions require a model to obtain the fair value.

Assume that the relationship between the true fair values of securities and the model-based estimates of these values is:

\[
\epsilon_i = c_i + c_i \epsilon_i
\]

where the estimation error \( \epsilon_i \) satisfies \( \epsilon_i \sim N(0, \sigma_i^2) \) for \( \sigma_i << 1 \) to keep \( \epsilon_i \) positive with very large probability, and \( \text{corr}(\epsilon_i, \epsilon_j) = \rho_{ij} \in [-1, 1] \). The assumption of zero mean for \( \epsilon_i \) reflects the bank’s belief that there is no systematic bias in the modeling of this security. Further assume that \( \epsilon_i \) is independent of \( c_i \), i.e., the estimation error for each security is independent of its price given by the model.

Assume that the financial institution's preference over the fair value of its holdings (\( \bar{V} \)) is:

\[
U(\bar{V}) = \bar{V} - \frac{k}{2} \bar{V}^2
\]

where \( k \) is a positive constant satisfying \( k < \frac{1}{\text{max} \bar{V}} \). The institution's expected utility of the fair value of its holdings is:

\[
E[U(\bar{V})] = E[\bar{V}] - \frac{k}{2} (E[\bar{V}]^2 + \text{VAR}[\bar{V}])
\]

where \( E[\bar{V}] \) and \( \text{VAR}[\bar{V}] \) are the mean and variance of \( \bar{V} \), respectively. This expected utility expression indicates that the bank is risk-averse regarding the uncertainty of the fair value of its
holdings.

Assume that the amount of cash held by the institution at time t is \( M_t \), which can be either positive, or negative, or zero. A negative cash position simply means that the bank has borrowed funds. The value of the institution’s holding at time t is:

\[
\bar{V}_t = M_t + \sum_{i=1}^{2} q_{i,t} c_{i,t} = M_t + \sum_{i=1}^{2} q_{i,t} c_{i,t}(1 + \epsilon_{i,t})
\]

The value of the same holding in the future---at time \( t + 1 \)---is:

\[
\bar{V}_{t+1} = M_t(1 + r_t) + \sum_{i=1}^{2} q_{i,t} c_{i,t+1} = M_t(1 + r_t) + \sum_{i=1}^{2} q_{i,t} c_{i,t+1}(1 + \epsilon_{i,t+1})
\]

where \( r_t \) is the value of the risk-free interest rate at t. The mean and variance of the fair value of the holdings at \( t + 1 \) are, respectively:

\[
\begin{align*}
E_{t+1} &\equiv E[\bar{V}_{t+1}] = M_t(1 + r_t) + \sum_{i=1}^{2} q_{i,t} E[c_{i,t+1}] = M_t(1 + r_t) + \sum_{i=1}^{2} q_{i,t} \bar{\mu}_{i,t+1} \\
\text{VAR}_{t+1} &\equiv \text{VAR}[\bar{V}_{t+1}] = \text{VAR}[\sum_{i=1}^{2} q_{i,t} c_{i,t+1}(1 + \epsilon_{i,t+1})]
\end{align*}
\]

\[
= \sum_{i=1}^{2} \sum_{j=1}^{2} q_{i,t} q_{j,t} \{ \bar{\sigma}_{i,t+1} \bar{\sigma}_{j,t+1} \bar{\rho}_{i,j} + (\bar{\sigma}_{i,t+1} \bar{\sigma}_{j,t+1} \bar{\rho}_{i,j} + \bar{\mu}_{i,t+1} \bar{\mu}_{j,t+1}) \sigma_i \sigma_j \rho_{i,j} \}
\]

where \( \bar{\mu}_{i,t+1} \) and \( \bar{\sigma}_{i,t+1} \) are the mean and standard deviation of \( c_{i,t+1} \), respectively, i.e., \( \bar{\mu}_{i,t+1} \equiv E[c_{i,t+1}] \), \( \bar{\sigma}_{i,t+1} \equiv \sqrt{\text{VAR}(c_{i,t+1})} \), and \( \bar{\rho}_{i,j} \) is the correlation between \( c_{i,t+1} \) and \( c_{j,t+1} \). The bank’s expected utility of the fair value of its holdings at \( t + 1 \) is:

\[
EU_{t+1} \equiv E_t[U(\bar{V}_{t+1})] = E_{t+1} - \frac{k}{2} [(E_{t+1})^2 + \text{VAR}_{t+1}]
\]

II.2 A Model Risk Metric

Because model risk (\( MR \)) affects only the variance of the fair value of the future holdings, and not the expected fair value, a natural choice for a metric on the extent of model risk is the increase in the variance of the fair value of the positions due to the presence of model risk. This is the difference between the variance of the fair value of the holdings under the existence of model risk and the variance when model risk is absent:

\[
MR_{t+1}(q_{1,t}, q_{2,t}) \equiv \text{VAR}_{t+1} - \text{VAR}_{t+1}\mid_{\sigma_1=\ldots=\sigma_7=0} = \sum_{i=1}^{2} \sum_{j=1}^{2} q_{i,t} q_{j,t} (\bar{\sigma}_{i,t+1} \bar{\sigma}_{j,t+1} \bar{\rho}_{i,j} + \bar{\mu}_{i,t+1} \bar{\mu}_{j,t+1}) \sigma_i \sigma_j \rho_{i,j}
\]

As can be seen in (2), the bank’s exposure to model risk depends on the values of three sets of variables. First, it depends on the institution’s holdings (\( q_{i,t} \) and \( q_{j,t} \)): (i) larger holdings are more likely to produce larger model risk and (ii) opposite holdings (long on one security and short on the other) are more likely to lead to reduction in model risk. Second, the exposure depends on the means and variances (including correlations) of the future values of the securities: (i) a larger mean (\( \bar{\mu}_{i,t+1} \) and \( \bar{\mu}_{j,t+1} \)) or a larger variance (\( \bar{\sigma}_{i,t+1} \) and \( \bar{\sigma}_{j,t+1} \)) is more likely to produce larger risk, (ii) higher positive (negative) correlations (\( \bar{\rho}_{i,j} \)) are more likely to produce higher (lower) risk. Third, it depends on both the sizes and correlations of the estimation errors: (i) larger estimation errors (\( \sigma_i \) and \( \sigma_j \)) are more likely to produce larger risk and (ii) positively (negatively) correlated estimation errors (\( \rho_{i,j} \)) are more likely to produce larger (smaller) risk.

II.3 Bid and Ask Prices and Model Risk Premia

We can derive the bid and ask prices and model-risk premia using the model risk model. Denote by \( b_{i,t}(\Delta q_{i,t}) > 0 \) and \( a_{i,t}(\Delta q_{i,t}) > 0 \) the bid and ask prices posted by the bank on the \( i \)th securities for a trade of size \( \Delta q_{i,t} \) on that security. These general results for the bid and ask prices
are formally derived in Appendix A.²

Proposition 1  (i) The bid price \( (b_{1,t}) \) that the institution would post for a trade of size \( \Delta q_{1,t} > 0 \) on the first security, which would not make the institution worse off, satisfies this condition:

\[
0 \leq \bar{\mu}_{1,t+1} - b_{1,t}(1 + r_t)
- \frac{k}{2} [\Delta q_{1,t} (\bar{\mu}_{1,t+1} - b_{1,t}(1 + r_t))^2 + 2(\bar{\mu}_{1,t+1} - b_{1,t}(1 + r_t))E_{t+1}]
- \frac{k}{2} (2q_{1,t} + \Delta q_{1,t})[\bar{\sigma}_{2,t+1}^2 + (\bar{\sigma}_{1,t+1}^2 + \bar{\mu}_{1,t+1}^2)\sigma_t^2]
- kq_{2,t} [\bar{\sigma}_{1,t+1}\bar{\sigma}_{2,t+1}\bar{\mu}_{1,t+1}\bar{\mu}_{2,t+1}^2 + (\bar{\sigma}_{1,t+1}\bar{\sigma}_{2,t+1}\bar{\mu}_{1,t+1}\bar{\mu}_{2,t+1}^2)\sigma_t^2\rho_{12}]
\]

(ii) The ask price \( (a_{1,t}) \) that the institution would post for a trade of size \( \Delta q_{1,t} < 0 \) on the first security, which would not make the institution worse off, satisfies:

\[
0 \geq \bar{\mu}_{1,t+1} - a_{1,t}(1 + r_t)
- \frac{k}{2} [\Delta q_{1,t} (\bar{\mu}_{1,t+1} - a_{1,t}(1 + r_t))^2 + 2(\bar{\mu}_{1,t+1} - a_{1,t}(1 + r_t))E_{t+1}]
- \frac{k}{2} (2q_{1,t} + \Delta q_{1,t})[\bar{\sigma}_{2,t+1}^2 + (\bar{\sigma}_{1,t+1}^2 + \bar{\mu}_{1,t+1}^2)\sigma_t^2]
- kq_{2,t} [\bar{\sigma}_{1,t+1}\bar{\sigma}_{2,t+1}\bar{\mu}_{1,t+1}\bar{\mu}_{2,t+1}^2 + (\bar{\sigma}_{1,t+1}\bar{\sigma}_{2,t+1}\bar{\mu}_{1,t+1}\bar{\mu}_{2,t+1}^2)\sigma_t^2\rho_{12}]
\]

The model risk premium is the additional amount on a trade with a customer that the bank deserves for bearing model risk. It can have two values for a given size of the trade. One is for increasing the bank’s holdings in the security--this is the premium adjusting the bid price. The other is for reducing the bank’s holdings in the security--this is the premium adjusting the ask price. We define the model risk premium on the \( i \)th security at the bid price to be the difference between the bid price when model risk does not exist and the bid price under the existence of model risk, and define the model risk premium at the ask price to be the difference between the ask price under the existence of model risk and the ask price when the model risk does not exist:

\[
\text{Prem}_{\text{bid}} (i) \equiv b_{i,t\mid\sigma_i=0} - b_{i,t\mid\sigma_i=0}
\]

\[
\text{Prem}_{\text{ask}} (i) \equiv a_{i,t\mid\sigma_i=0} - a_{i,t\mid\sigma_i=0}
\]

The model risk premia are defined in this manner to conform with the intuition that a market maker will be expected to widen its bid-ask spread when encountering greater risk. The bid price is lowered and the ask price is raised to compensate the market maker for bearing that risk, and the risk premia as defined measures that compensation. Our model allows the bank to have short as well as long positions, and to set its bid and ask prices on trades with customers to increase or decrease its inventory depending on expected futures prices on the securities. In general, the risk premia are likely to be affected by four sets of variables: the institution’s current security holdings, the mean and variance of the future values of securities, sizes and correlations of the estimation errors, and the size of the trade. An important element to the model is that the bank makes its pricing decisions based not only on the profit from an individual trade, but also weighs the return over time on its balance sheet assets and liabilities that are impacted by that trade. The presence of model risk always leads to a wider bid-ask spread than otherwise but, depending on the parameters, can also lead to a change in the level of the bid and ask prices. In that circumstance one of the model risk premia can be negative.

II.4 Special Case of Symmetry

To keep the number of parameters that we have to work with reasonably small, we next consider a special case of the model---the case of symmetry whereby the two securities have the same

²Here we provide only the results for the first security. Those for the second security can be obtained similarly.
When computing the prices for the first security, we assume that a trade would occur on this security and no trade would occur on the other security.
characteristics and the bank has identical holdings of each. We make the following simplifying assumptions on the parameters and variables in the model:

\[ c_{1,t} = c_{2,t} = c \]
\[ \bar{\mu}_{1,t+1} = \bar{\mu}_{2,t+1} = \bar{\mu} = c(1 + r)(1 + \hat{\mu}) \]
\[ \bar{\sigma}_{1,t+1} = \bar{\sigma}_{2,t+1} = c\hat{\sigma} \]
\[ \sigma_{1,t+1} = \sigma_{2,t+1} = \sigma \]
\[ q_{1,t} = q_1 \]
\[ q_{2,t} = q_2 \]
\[ \Delta q_{1,t} = \Delta q_1 \]
\[ b_{1,t} = b_1 \]
\[ a_{1,t} = a_1 \]
\[ M_t = M \]
\[ r_t = r \]

In (6), the first equality indicates that the two securities have the same estimated values at time \( t \) based on the models. \( \hat{\mu} \) in the second equality is (approximately) the excess return of the security over the risk-free rate. \( \hat{\sigma} \) in the third equality is the volatility of returns on the securities.

Under the assumptions in (6), we derive in Appendix B the following results for the bid and ask prices for the first security.

**Proposition 2** (i) The bid price \( (b_{1,t}) \) for a trade of size \( \Delta q_{1,t} > 0 \) on the first security which would not make the institution worse off satisfies this condition:

\[ b_{1,t} \in [\underline{b}_1, \bar{b}_1] \]

where

\[
\underline{b}_1 = c(1 + \hat{\mu}) + \frac{b_B}{A_B} - \frac{c_B}{B_B}
\]
\[
\bar{b}_1 = c(1 + \hat{\mu}) + \frac{c_B}{B_B}
\]

and

\[
A_b = \frac{k}{2} \Delta q_{1}^b (1 + r)^2
\]
\[
B_b = k(M + c(q_1 + q_2))(1 + r)^2 - (1 + r)
\]
\[
C_b = \frac{k}{2} (2q_1 + \Delta q_{1}^b)c^2[\hat{\sigma}^2 + (\hat{\sigma}^2 + (1 + r)^2)\sigma^2]
\]
\[
- kq_2 c^2[\hat{\sigma}^2 \bar{\rho}_{12} + (\hat{\sigma}^2 \bar{\rho}_{12} + (1 + r)^2)\sigma^2 \rho_{12}]
\]

(ii) The ask price \( (a_{1,t}) \) for a trade of size \( \Delta q_{1,t} < 0 \) on the first security which would not make the institution worse off satisfies:

\[ a_1 \in [\underline{a}_1, \bar{a}_1] \]

where

\[
\underline{a}_1 = c(1 + \hat{\mu}) + \frac{c_B}{B_a}
\]
\[
\bar{a}_1 = c(1 + \hat{\mu}) + \frac{b_B}{A_a} - \frac{c_B}{B_a}
\]

and

\[
A_a = -\frac{k}{2} \Delta q_{1}^a (1 + r)^2
\]
\[
B_a = 1 + r - k(M + c(q_1 + q_2))(1 + r)^2 = -B_b
\]
\[
C_a = -\frac{k}{2} (2q_1 + \Delta q_{1}^a)c^2[\hat{\sigma}^2 + (\hat{\sigma}^2 + (1 + r)^2)\sigma^2]
\]
\[
- kq_2 c^2[\hat{\sigma}^2 \bar{\rho}_{12} + (\hat{\sigma}^2 \bar{\rho}_{12} + (1 + r)^2)\sigma^2 \rho_{12}]
\]
II.5 Effects of Model Risk on the Pricing and Trading of the Securities

Next we compute the model risk premia and discuss the effects of model risk on the pricing and trading of the two illiquid securities. We will base the observations on the results for the highest bid price and the lowest ask price given in the previous section.

Define the smallest bid-ask spread ($\Delta$) to be the difference between the lowest ask price and the highest bid price. Using the results above, we can solve for these prices:

$$
\bar{b}_1 = c + c\mu + \frac{c_1-c_2\Delta q_1^b}{B_a} \\
\underline{a}_1 = c + c\mu + \frac{c_1+c_2|\Delta q_1^q|}{B_a} \\
\Delta \equiv \underline{a}_1 - \bar{b}_1 = \frac{c_2(\Delta q_1^b+|\Delta q_1^q|)}{B_a}
$$

where

$$B_a = (1+r)-(M+c(q_1+q_2))(1+r)^2 > 0$$

$$C_1 = -kq_1c^2[\sigma^2+(\sigma^2+(1+r)^2)\sigma^2]$$

$$\quad- kq_2c^2[\sigma^2\tilde{\rho}_{12}+(\sigma^2\tilde{\rho}_{12}+(1+r)^2)\sigma^2\rho_{12}]$$

$$C_2 = \frac{k}{2}c^2[\sigma^2+(\sigma^2+(1+r)^2)\sigma^2]$$

Following the definitions stated in (5), the risk premia for bid and ask prices can be computed to be:

$$
Prem_{bid} \equiv \bar{b}_1|_{\sigma=0} - \bar{b}_1|_{\sigma \neq 0} \\
= \frac{kq_1c^2\sigma^2}{B_a}(\bar{\sigma}^2+(1+r)^2) + \frac{kq_2c^2\sigma^2\rho_{12}}{B_a}(\bar{\sigma}^2\tilde{\rho}_{12}+(1+r)^2) \\
+ \frac{kc^2\sigma^2}{2B_a}\Delta q_1^b(\sigma^2+(1+r)^2) \\
= \frac{k}{2B_a}MR_{t+1}(q_1+\Delta q_1^b,q_2)-MR_{t+1}(q_1,q_2)
$$

$$
Prem_{ask} \equiv \underline{a}_1|_{\sigma=0} - \underline{a}_1|_{\sigma \neq 0} \\
= -\frac{kq_1c^2\sigma^2}{B_a}(\bar{\sigma}^2+(1+r)^2) - \frac{kq_2c^2\sigma^2\rho_{12}}{B_a}(\bar{\sigma}^2\tilde{\rho}_{12}+(1+r)^2) \\
+ \frac{kc^2\sigma^2}{2B_a}|\Delta q_1^q|((\sigma^2+(1+r)^2) \\
= -\frac{k}{2B_a}MR_{t+1}(q_1-|\Delta q_1^q|,q_2)-MR_{t+1}(q_1,q_2)
$$

Several conclusions can be drawn from these results. First, as shown in (9) to (11), the bid and ask prices, the bid-ask spread, and the model-risk premia depend on four sets of variables: (i) the institution’s current holdings of the securities ($q_1$ and $q_2$), (ii) the mean and variance of future values of securities ($\mu$ and $\sigma$), (iii) variance of estimation errors from the model ($\sigma$), and (iv) size of the trade ($\Delta q_1^b$ and $\Delta q_1^q$). Since these variables affect both the mean and variance of the future fair value of the bank’s holdings, they affect the model-risk premia.

Second, higher (lower) expected security prices would lead to both higher (lower) bid and ask prices because both prices are increasing in the expected future values of the securities ($c\mu$). The bank, while profiting on each trade because the bid-ask spread is always positive ($\Delta > 0$), also potentially benefits from higher expected returns on the securities held in the trading portfolio. In that case, the financial institution is willing to pay a higher price to get more inventory and will only want to sell if a higher sale price is received. Lower expected prices motivate the bank to reduce inventory and perhaps even move to a short holding in the security. That is accomplished by lowering the ask price to provide an incentive for customers to buy and also lowering the bid price to make customers more reluctant to sell.
Third, other things being equal, current long (short) positions in the securities would lead to lower (higher) bid and ask prices. Notice that both bid and ask prices are decreasing in the institution's holding on the first security \( q_1 \). When the bank already has a reasonably large long (short) position on the first security, the institution may find it optimal to decrease (increase) its bid and ask prices to reduce (add to) inventory. The bank gives investors more (less) incentive to purchase the security from the bank and less (more) incentive to sell it to the bank.

Fourth, due to the effects of expected future values of securities and the institution's current holdings on bid and ask prices discussed above, we can observe mis-pricing, i.e., either intentional over-pricing of the securities (which occurs when \( c < \tilde{b} < a \)) or under-pricing of the securities (which occurs when \( \tilde{b} < a < c \)) or fair pricing of the securities (which occurs when \( \tilde{b} < c < a \)). Over-pricing and under-pricing mean that the bid and ask prices are above (below) those indicated by the model. According to (9), a low expected future value for the security and reasonably large long positions are more likely to lead to under-pricing of the securities. This is because both bid and ask prices are increasing in the expected future values \( c\hat{\mu} \) and in decreasing in the institution's holdings on the securities \( q_1 \). High expected future value and reasonably large short positions on the securities are more likely to lead to over-pricing for the same reason. Medium expected future values and small holdings on the securities are less likely to lead to mis-pricing of the securities.

Fifth, model risk can have a big impact on the bid and ask prices. When the bank has long (short) positions on both securities, model risk would further decrease (increase) the bid and ask prices, which would make the mis-pricing more severe. To see this result, notice that according to (9) to (11), when \( q_1 \) and \( q_2 \) are positive (negative), larger values of \( \sigma \) would further reduce (raise) the values of the bid and ask prices. The intuition behind this result is that model risk adds additional uncertainty about the future values of the assets, which would make the institution behave as if it is more risk averse. When the bank has a reasonably large long (short) position on the first security, model risk would lead to further deductions in (additions to) the bid and ask prices, which would give customers even more (less) incentive to buy the asset from the bank and less (more) incentive to sell it to the bank. An increase in the size of the trade would further decrease (increase) the bid price due to the bank's reluctance to have a larger long (short) position on the asset, but would increase (decrease) the ask price.

Sixth, model risk affects securities with different price volatilities differently. It affects securities with low volatility of returns much more than it does to securities with high volatility, because for securities with low (high) volatility, model risk (market risk) is the main risk facing the market maker. Too see this result, notice that in (10), when the price volatility \( \hat{\sigma} \) is large (small) relative to the model risk \( \tilde{\sigma} \), the mis-pricing (given respectively by the last terms in the bid and ask prices) mainly depends on \( \hat{\sigma} \) vis-a-vis \( \tilde{\sigma} \).

Seventh, as the volatility of asset returns and/or the volatility of pricing errors increases, the extent of over-pricing and under-pricing, the bid-ask spread, and the model risk premia all increase because they are increasing in these variables. Too see this result, notice that all these three terms, i.e., mis-pricing (which is equal to the absolute values of the last terms in the bid and ask prices), bid-ask spread \( \Delta \), and risk premia are all increasing in \( \tilde{\sigma} \) and \( \sigma \).

Last, the holding on the second security can affect the bid and ask prices, the bid-ask spread, and the model risk premia of the first security only through the correlation of asset returns \( \tilde{\rho}_{12} \) and correlation of pricing errors \( \rho_{12} \). This is because, according to (9) to (11), the variable \( q_2 \) appears together with \( \tilde{\rho}_{12} \) and \( \rho_{12} \) in these three variables.
II.6 Numerical Examples

Some numerical examples are useful to illustrate some of the comparative static properties of the model. The results for the bid and ask prices, the bid-ask spread, and the bid and ask model risk premia for six scenarios are reported in Table 1. The calculations for the third scenario are shown in Appendix C; the others obtained in the same manner. The statistical model is assumed to produce a value of $10 for each of the securities ($EgR_{μ}55 Eg\sum_{i=4}^{A} 10$). The problem is to set the bid and ask prices for trade sizes of one unit ($EgR_{μ}9\equiv EgR_{μ}9\equiv 9\equiv EgM_{μ}9\equiv 9\equiv Eg\sum_{i=4}^{A} 1$) given an assumed interest rate ($EgR_{μ}7A Eg\sum_{i=4}^{A} .05$) and level of market risk ($EgM_{μ}5 Eg\sum_{i=4}^{A} 0.10$). To keep the examples simple, we assume no correlation in the security returns ($EgM_{μ}5 Eg\sum_{i=4}^{A} 0$) or in the estimation errors ($EgM_{μ}5 Eg\sum_{i=4}^{A} 0$). Also, we will assume that the key variable in the expected utility function allows for reasonable results ($EgR_{μ}\equiv \sum_{i=4}^{A} 0.005 Eg\equiv \equiv 7.13$).

Given these parameter values, we consider in the examples changes in the bank's current balance sheet ($EgR_{μ}\equiv 9\equiv EgM_{μ}\equiv 9\equiv EgR_{μ}\equiv 9\equiv EgM_{μ}7A$), the presence of model risk ($EgM_{μ}\equiv 0$), and expected returns on the securities ($EgM_{μ}\equiv 0$).

Scenario (1) provides a base case where there is no model risk ($EgM_{μ}\equiv 0$), meaning that the statistical model produces the true value for the illiquid securities, and no expected change in the value of the securities ($\mu = 0$). The bank is currently long in both securities, holding 10 units of each ($q_1 = q_2 = +10$), as typically would be the case for a market maker. Presumably, these positions have been acquired with borrowed funds ($M = -100$). The bid price would be set at 9.8947 and the ask price at 9.9048, so that the bid-ask spread is 0.0101. Notice that there is under-pricing because both bid and ask prices are below the value produced by the model ($EgR_{μ}55 Eg\sum_{i=4}^{A} 10$). This is because of risk aversion. The bank will set the bid and ask prices to reduce risk by reducing inventory. Under-pricing encourages customers to buy at the bank's lowered ask price and discourages them from selling securities to the bank at its lowered bid price. This is a standard result in the market microstructure literature.

Scenario (2) adds model risk ($EgM_{μ}\equiv 0.10$) to the base case example. The presence of model risk exacerbates the under-pricing in that the bid price is lowered to 9.7776 and the ask price to 9.7988. In this case the model risk premium to the bid price is positive, but it is negative to the ask price. The combined effect though is to widen the bid-ask spread to 0.0212, as would be expected given an additional risk factor. This exemplifies our main point in this paper---model risk can be an important determinant of the bid-ask spread.

Scenarios (3) and (4) show the impact of positive and negative expected returns on the securities. In scenario (3) when the securities are expected to appreciate in value ($\mu = +0.08$), the bank will want to increase its inventory. Both prices are raised---the bid price goes up to 10.5776 and the ask price to 10.5988 but the spread remains at 0.0212 to compensate the bank for both market risk and model risk. Even with long positions in the securities, there is over-pricing. In scenario (4) the bank will lower its prices to reduce its holdings because the security is expected to depreciate in value ($\mu = -0.08$). Relative to scenario (2), there is greater under-pricing. The bid price is lowered to 8.9776 and the ask price to 8.9988, providing the same spread of 0.0212.

Scenario (5) establishes the base case for the impact of short positions in both securities ($q_1 = q_2 = -10$) when there is again no model risk ($EgM_{μ}\equiv 0$) and no expected appreciation or depreciation in the value of the securities ($\mu = 0$). We assume that the bank has borrowed the securities from another bank and sold them to customers to meet their demand, thereby generating a positive cash position ($M = +300$). This is another example of over-pricing because both prices are greater than the value obtained from the model ($EgR_{μ}\equiv 10$). The bid price is

---

3 Of the $300, $200 comes from selling the two securities. $100 is the cash that the bank has.
10.0952 and the ask price is 10.1053. Being risk averse, the bank would like to reduce its short positions—the higher bid price encourages customers to sell the security and the higher ask price discourages them from wanting to buy more from the bank. As in scenario (1), the bid-ask spread is 0.0101.

Scenario (6) shows that model risk \((E_{GAM} \equiv \sum_{A4} 0.10)\) reinforces the over-pricing when the bank has short positions in the securities. Compared to scenario (5), the bid price goes up to 10.2012 and the ask price to 10.2224. The bid-ask spread is widened to 0.0212 to compensate the bank for the presence of model risk. In this case, the bid model risk premium is negative and the ask model risk premium is positive. This is the opposite to the result in scenario (2), but once again the bid-ask spread is widened by the same amount to compensate the financial institution for both market risk and model risk.\(^4\)

### III. Conclusions

This paper studies the effects of model risk on the pricing and trading of illiquid securities. The analysis shows that setting the bid and ask prices when a bank makes a two-way market can lead to intentional over-pricing and under-pricing of securities vis-a-vis the fair values estimated using the model. This is because the bank considers the return on its inventory along with the profitability of each trade. In some circumstances, the financial institution will under-price, thereby lowering both its bid and ask prices, in order to reduce long holdings and perhaps even take on a short position. In other circumstances, the bank might over-price to build up its trading inventory. A significant part of this over-pricing and under-pricing can be due to model risk. In general, the bid and ask prices, bid-ask spread, and model risk premia depend on the values of four sets of variables—the institution's holdings of securities, the mean and variance of the future returns of securities, the sizes and correlations of the estimation errors, and the size of the transaction.

We believe that model risk is an important factor in market microstructure. One must keep in mind, however, that all of these results have been produced by yet another model. Inevitably, there is model risk to a model risk model.

### References


Hull, J., Suo, W., 2002, "A methodology for assessing model risk and its applications to the

\(^4\)According to (11), when we choose the security holdings in Scenarios (5) and (6) to be the opposite of those in Scenarios (1) and (2) while keeping the bank's wealth unchanged, the bid (ask) model risk premia in Scenarios (5) and (6) become the ask (bid) model risk premia in Scenarios (1) and (2).
Appendix A

In this section we only compute the bid and ask prices for the first security. Those for the second security can be computed similarly. After a trade of size $\Delta q_{1,t}$ on the first security at $t$, the fair value of the bank's holding at $t + 1$ is:

$$\bar{V}_{t+1} = M_{t+1} + (q_{1,t} + \Delta q_{1,t})c_{1,t+1}(1 + \varepsilon_{1,t+1}) + q_{2,t}c_{2,t+1}(1 + \varepsilon_{2,t+1})$$

where

$$M_{t+1} = \begin{cases} 
(M_t - b_{1,t}\Delta q_{1,t})(1 + r_t) & \Delta q_{1,t} > 0 \\
(M_t - a_{1,t}\Delta q_{1,t})(1 + r_t) & \Delta q_{1,t} < 0 
\end{cases}$$

The mean and variance of the future fair value of holdings are:

$$E_{t+1}' = E[\bar{V}_{t+1}] = M_{t+1} + (q_{1,t} + \Delta q_{1,t})\mu_{1,t+1} + q_{2,t}\mu_{2,t+1}$$

$$VAR_{t+1}' = VAR[\bar{V}_{t+1}] = (q_{1,t} + \Delta q_{1,t})^2[\sigma_{1,t+1}^2 + (\sigma_{1,t+1}^2 + \mu_{1,t+1}^2)\sigma_{1,t+1}^2]$$

$$+ q_{2,t}^2[q_{2,t}[(\bar{\sigma}_{1,t+1}\bar{\sigma}_{2,t+1}\bar{\mu}_{1,2} + (\bar{\sigma}_{1,t+1}\bar{\sigma}_{2,t+1}\bar{\mu}_{1,2} + \bar{\mu}_{1,t+1}\bar{\mu}_{2,t+1})\sigma_1\sigma_2\rho_{12}]]$$

The bank's expected utility of the future fair value of the holdings is:

$$EU_{t+1} = E[\bar{V}_{t+1}] = E_{t+1}' - \frac{k}{2}[(E_{t+1}')^2 + VAR_{t+1}']$$

For the bid price, equation (12) can be written as follows using equation (1):

$$E_{t+1}' = E_{t+1} + \Delta q_{1,t}(\bar{\mu}_{t+1} - q_{1,t} + r_t))$$

$$VAR_{t+1}' = VAR_{t+1} + (2q_{1,t}\Delta q_{1,t} + (\Delta q_{1,t})^2)[\bar{\sigma}_{1,t+1}^2 + (\bar{\sigma}_{1,t+1}^2 + \bar{\mu}_{1,t+1}^2)\sigma_1^2]$$

$$+ 2\Delta q_{1,t}q_{2,t}[\bar{\sigma}_{1,t+1}\bar{\sigma}_{2,t+1}\bar{\mu}_{1,2} + (\bar{\sigma}_{1,t+1}\bar{\sigma}_{2,t+1}\bar{\mu}_{1,2} + \bar{\mu}_{1,t+1}\bar{\mu}_{2,t+1})\sigma_1\sigma_2\rho_{12}])$$

The bid price that the bank would post should satisfy the condition that a trade in which the bank buys the first security in the amount of $\Delta q_{1,t}$ should not make the bank worse off. In other words, it should not decrease the bank's expected utility of future fair value of its holdings. This condition gives us, for $\Delta q_{1,t} > 0$:

$$0 \leq \Delta EU = EU_{t+1} - EU_{t+1}$$

$$= \Delta q_{1,t}(\bar{\mu}_{t+1} - q_{1,t} + r_t))$$

$$- \frac{k}{2}[(\Delta q_{1,t})^2(\bar{\mu}_{t+1} - q_{1,t} + r_t))^2 + 2\Delta q_{1,t}(\bar{\mu}_{t+1} - q_{1,t} + r_t))E_t)]$$

Simplifying this inequality gives us another inequality on $b_{1,t}$, from which the range of the bid price can be determined:

$$0 \leq \bar{\mu}_{t+1} - q_{1,t} + r_t$$

$$- \frac{k}{2}[(\Delta q_{1,t}(\bar{\mu}_{t+1} - q_{1,t} + r_t))^2 + 2(\bar{\mu}_{t+1} - q_{1,t} + r_t))E_t)]$$

Similarly, we can show that the ask price at which the bank sells the first security satisfies this condition for $\Delta q_{1,t} < 0$:

$$0 \geq \bar{\mu}_{t+1} - q_{1,t} + r_t$$

$$- \frac{k}{2}[(\Delta q_{1,t}(\bar{\mu}_{t+1} - q_{1,t} + r_t))^2 + 2(\bar{\mu}_{t+1} - q_{1,t} + r_t))E_t)]$$

716
Appendix B

This section derives the bid and ask prices of the first security in the symmetric case. Equation (1) and equation (6) together are:

\[ E_{t+1} = M(1 + r) + \bar{\mu}(q_1 + q_2) = [M + c(1 + \bar{\mu})(q_1 + q_2)](1 + r) \]

According to equation (3), for \( \Delta q_1^b > 0 \) the bid price satisfies:

\[ 0 \leq (c(1 + \bar{\mu}) - b_1)(1 + r) \]

\[ -\frac{k}{2} [\Delta q_1^b(c(1 + \bar{\mu}) - b_1)^2(1 + r)^2 + 2(c(1 + \bar{\mu}) - b_1)(1 + r)E_{t+1}] \]

\[ -\frac{k}{2} (2q_1 + \Delta q_1^b)c^2[\bar{\sigma}^2 + (\bar{\sigma}^2 + (1 + r)^2)\sigma^2] \]

\[ -kq_2c^2[\bar{\sigma}^2\bar{\rho}_{12} + (\bar{\sigma}^2\bar{\rho}_{12} + (1 + r)^2)\sigma^2\rho_{12}] \]

This can be written more compactly as:

\[ A_b\Delta_b^2 + B_b\Delta_b + C_b \leq 0 \]

where:

\[ \Delta_b = c(1 + \bar{\mu}) - b_1 \]

\[ A_b = \frac{k}{2} \Delta q_1^b (1 + r)^2 \]

\[ B_b = k(M + c(q_1 + q_2))(1 + r)^2 - (1 + r) \]

\[ C_b = \frac{k}{2} (2q_1 + \Delta q_1^b)c^2[\bar{\sigma}^2 + (\bar{\sigma}^2 + (1 + r)^2)\sigma^2] \]

\[ +kq_2c^2[\bar{\sigma}^2\bar{\rho}_{12} + (\bar{\sigma}^2\bar{\rho}_{12} + (1 + r)^2)\sigma^2\rho_{12}] \]

\[ \Delta q_1^b > 0 \]

Solving for this inequality gives us the range of the bid price:

\[ b_1 \in [\bar{b}_1, \bar{b}_1] \]

where:

\[ b_1 = c(1 + \bar{\mu}) + \frac{B_b - \sqrt{B_b^2 - 4A_bC_b}}{2A_b} \]

\[ \bar{b}_1 = c(1 + \bar{\mu}) + \frac{B_b + \sqrt{B_b^2 - 4A_bC_b}}{2A_b} \]

Equation (13) produces:

\[ b_1 = c(1 + \bar{\mu}) + \frac{B_b - \sqrt{B_b^2 - 4A_bC_b}}{2A_b} \]

\[ \bar{b}_1 = c(1 + \bar{\mu}) + \frac{B_b + \sqrt{B_b^2 - 4A_bC_b}}{2A_b} \]

For small enough \( \Delta q_1^b \), \( \sqrt{B_b^2 - 4A_bC_b} \) can be approximated by:

\[ \sqrt{B_b^2 - 4A_bC_b} = -B_b \left( 1 - \frac{4A_bC_b}{B_b^2} \right)^{1/2} \approx -B_b \left( 1 - \frac{4A_bC_b}{2B_b^2} \right) \]

since \( B_b < 0 \). Substituting this approximation into equation (13) produces:

\[ b_1 = c(1 + \bar{\mu}) + \frac{B_b}{A_b} - \frac{C_b}{B_b} \]

\[ \bar{b}_1 = c(1 + \bar{\mu}) + \frac{B_b}{A_b} + \frac{C_b}{B_b} \]

According to equation (4), for \( \Delta q_1^a < 0 \) the ask price satisfies this condition:

\[ 0 \geq (c(1 + \bar{\mu}) - a_1)(1 + r) \]

\[ -\frac{k}{2} [\Delta q_1^a(c(1 + \bar{\mu}) - a_1)^2(1 + r)^2 + 2(c(1 + \bar{\mu}) - a_1)(1 + r)E_{t+1}] \]

\[ -\frac{k}{2} (2q_1 + \Delta q_1^a)c^2[\bar{\sigma}^2 + (\bar{\sigma}^2 + (1 + r)^2)\sigma^2] \]

\[ -kq_2c^2[\bar{\sigma}^2\bar{\rho}_{12} + (\bar{\sigma}^2\bar{\rho}_{12} + (1 + r)^2)\sigma^2\rho_{12}] \]

This can be re-stated:

\[ A_a\Delta_a^2 + B_a\Delta_a + C_a \leq 0 \]

where:

\[ \Delta_a = c - a_1 \]
\[ A_a = -\frac{k}{2} \Delta q_a^2 (1 + r)^2 \]
\[ B_a = 1 + r - k(M + c(q_1 + q_2))(1 + r)^2 = -B_b \]
\[ C_a = -\frac{k}{2} (2q_1 + \Delta q_a^2)c^2[\hat{\sigma}^2 + (\hat{\sigma}^2 + (1 + r)^2)\hat{\sigma}^2] - kq_2c^2[\hat{\sigma}^2 \hat{\rho}_{12} + (\hat{\sigma}^2 \hat{\rho}_{12} + (1 + r)^2)\sigma_0] \]
\[ \Delta q_a^2 < 0 \]

Solving for this inequality gives us the range of the ask price:
\[ a_1 \in [\bar{a}_1, \bar{a}_1] \]

where:
\[ a_1 = c(1 + \hat{\mu}) + \frac{B_a - \sqrt{B_a^2 - 4A_aC_a}}{2A_a} \quad (14) \]
\[ \bar{a}_1 = c(1 + \hat{\mu}) + \frac{B_a + \sqrt{B_a^2 - 4A_aC_a}}{2A_a} \]

For small enough \( \Delta q_a^2, \sqrt{B_a^2 - 4A_aC_a} \) can be approximated by:
\[ \sqrt{B_a^2 - 4A_aC_a} = B_a \left( 1 - \frac{4A_aC_a}{B_a^2} \right)^{1/2} \approx B_a \left( 1 - \frac{4A_aC_a}{2B_a} \right) \]
since \( B_a > 0 \). Substituting this approximation into equation (14) produces:
\[ a_1 = c(1 + \hat{\mu}) + \frac{C_a}{B_a} \]
\[ \bar{a}_1 = c(1 + \hat{\mu}) + \frac{B_a - C_a}{B_a} \]

**Appendix C**

**Calculation of the Numerical Results in Scenario (3)**

In each of the scenarios, we assume that \( c = 10, \Delta q_1 = 1, \Delta q_2 = -1, r = .05, \hat{\sigma} = .10, \hat{\rho}_{12} = 0, \rho_{12} = 0, \) and \( k = .005 \). In Scenario (3) we assume that \( q_1 = q_2 = +10, M = -100, \sigma = .10, \) and \( \hat{\mu} = .08 \). From equations (9) to (11), we have:
\[ B_a = (1 + r) - k(M + c(q_1 + q_2))(1 + r)^2 = 1.05 - [0.005 * (-100 + 10 * (10 + 10)) * 1.05^2] = 0.498750 \]
\[ C_1 = -kq_1c^2[\hat{\sigma}^2 + (\hat{\sigma}^2 + (1 + r)^2)\hat{\sigma}^2] = -0.005 * 10 * 10^2 * 0.10^2 + (0.10^2 + 1.05^2) * 0.10^2 \]
\[ = -0.105625 \]
\[ C_2 = \frac{k}{2} c^2[\hat{\sigma}^2 + (\hat{\sigma}^2 + (1 + r)^2)\sigma^2] = 0.005 * 10^2 * 0.10^2 + (0.10^2 + 1.05^2) * 0.10^2 \]
\[ = 0.00528125 \]

and
\[ \bar{b}_1 = c(1 + \hat{\mu}) + \frac{C_1 - C_2\Delta q_1^2}{B_a} \]
\[ = (10 * 1.08) + \frac{-0.105625 - 0.00528125 * 1}{0.498750} = 10.57763 \]

\[ a_1 = c(1 + \hat{\mu}) + \frac{C_1 + C_2\Delta q_1^2}{B_a} \]
\[ = (10 * 1.08) + \frac{-0.105625 + 0.00528125 * 1}{0.498750} = 10.59881 \]
\[ \Delta \equiv a_1 - \bar{b}_1 = \frac{C_2(\Delta q_1^2 + |\Delta q_1^2|)}{B_a} = \frac{0.00528125 * 2}{0.498750} = 0.02118 \]
\[
\text{Prem}_\text{bid} = \frac{kq_1c^2\sigma^2}{B_a}(\hat{\sigma}^2 + (1 + r)^2) + \frac{k\sigma^2}{2B_a} \Delta q_b(\hat{\sigma}^2 + (1 + r)^2)
\]
\[
= \frac{0.005 \times 10^2 \times 0.10^2}{0.498750} \times (0.10^2 + 1.05^2)
+ \frac{0.005 \times 10^2 \times 0.10^2}{2 \times 0.498750} \times 1 \times (0.10^2 + 1.05^2)
= 0.11153 + 0.00558 = 0.11711
\]

\[
\text{Prem}_\text{ask} = -\frac{kq_1c^2\sigma^2}{B_a}(\hat{\sigma}^2 + (1 + r)^2) + \frac{k\sigma^2}{2B_a} |\Delta q_b| (\hat{\sigma}^2 + (1 + r)^2)
\]
\[
= -\frac{0.005 \times 10^2 \times 0.10^2}{0.498750} \times (0.10^2 + 1.05^2)
+ \frac{0.005 \times 10^2 \times 0.10^2}{2 \times 0.498750} \times 1 \times (0.10^2 + 1.05^2)
= -0.11153 + 0.00558 = -0.10595
\]
Table 1: Numerical Examples

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Bid Price</th>
<th>Ask Price</th>
<th>Bid-Ask Spread</th>
<th>Bid Model Risk Premium</th>
<th>Ask Model Risk Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_1 = q_2 = +10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M = -100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu} = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma} = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.8947</td>
<td>9.9048</td>
<td>0.0101</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_1 = q_2 = +10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M = -100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu} = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma} = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.7776</td>
<td>9.7988</td>
<td>0.0212</td>
<td>0.1171</td>
<td>-0.1060</td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_1 = q_2 = +10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M = -100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu} = 0.08$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma} = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.5776</td>
<td>10.5988</td>
<td>0.0212</td>
<td>0.1171</td>
<td>-0.1060</td>
</tr>
<tr>
<td>(4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_1 = q_2 = +10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M = -100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu} = -0.08$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma} = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.9776</td>
<td>8.9988</td>
<td>0.0212</td>
<td>0.1171</td>
<td>-0.1060</td>
</tr>
<tr>
<td>(5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_1 = q_2 = -10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M = +300$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu} = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma} = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.0952</td>
<td>10.1053</td>
<td>0.0101</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_1 = q_2 = -10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M = +300$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu} = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma} = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma = 0.10$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.2012</td>
<td>10.2224</td>
<td>0.0212</td>
<td>-0.1060</td>
<td>0.1171</td>
</tr>
</tbody>
</table>
Managing Credit Risk - a Method for Maintaining an Optimal Distance to Default

Jana Simon, PhD
Assistant Professor of Finance
Bittner School of Business
St. John Fisher College
Rochester, NY, 14618
jsimon@sjfc.edu
Abstract
This paper proposes a credit risk management method for targeting an optimal distance to default. It introduces a credit availability spread that governs an adaptable threshold of default. The changing aspects of the default threshold depend on system volatility and system liquidity, as it integrates likelihood of system-wide credit line disruptions. In order to achieve a steady distance I dynamically adjust asset portfolio volatilities.

JEL Classification: C2, G32, C54.

I. Introduction
This paper proposes a method for dynamic credit risk management. Firstly, I outline a credit accessibility spread S reflecting the easiness with which a firm can secure an on-going financing. Employing S enables me to define an elastic default boundary. Secondly I measure a firm’s distance to default given its threshold of default (ToD). Consequently I adjust a firm’s asset portfolio volatility to arrive at a target distance to default (DtD). The optimal distance to default is characterized as such asset volatility that allows firms to maximize their asset returns while minimizing credit risk.

Beginning with Merton (1974) and Black & Cox (1976), most structural models of credit risk assume that the firm defaults when the market value of its assets falls below some threshold reflecting debt servicing commitments. In this framework, any temporary cash flow shortfall can be alleviated by raising external financing – as long as the value of assets remains well above the default threshold. Vasicek (1984) states that if the value of assets is less than the amount of debt, which is not yet due, the firm can and will continue operating. Crosbie, Bohn (2003) conclude similarly that firms do not default when their asset value reaches the book value of their total liabilities. Consequently, firms’ asset values are allowed to decline to zero prior to its liabilities’ maturity. This paper proposes a method of active monitoring and managing of credit risk exposure.

I use ToD as a benchmark for determining the firms’ default probability prior to its liabilities’ maturity. ToD quantifies the extent to which the value of assets shall be allowed to fluctuate relative to the amount of debt, which is not yet due. A firm’s ToD is a function of the firm’s cost of capital, time to maturity, and credit accessibility. As cost of capital and time to maturity are fully exogenous. I model credit accessibility spread S as a function of a firm’s riskiness measured by beta, market volatility measured by the VIX® index, and credit line availability given by yield curve data.

There is a growing body of evidence that default risk and liquidity risk are correlated. Das et al. (2006) state that when economy-wide volatility is high, correlations between individual firms’ volatilities increase. They add that an inverted yield curve is a likely precursor of system-wide stress. Thus we observe ToD contracting with high market volatility and an inverted yield curve. Both of these factors have an undesirable effect on financing availability and so they act restrictively in spread S.

Credit risk management involves measuring the distance between firms’ asset values and their ToD. We can express DtD as a number of standard deviations from the ToD. This makes the DtD a standardized measure for companies of different sizes and/or liabilities positions. We can

5 The CBOE Volatility Index® (VIX®) is a key measure of market expectations of near-term volatility conveyed by S&P500 stock index option prices.

target a desirable DtD, as well as adjust our asset holdings and corresponding volatility to
achieve the target. “In the presence of financing constraints and costs of financial distress,
riskier firms may choose to maintain higher cash reserves in order to reduce the possibility of a
cash shortage in the future. The firm optimally increases its cash holdings, which reduces
probability of a cash shortfall when the debt comes due.” Acharya, et al. (2008). A steady DtD
thus requires an elastic ToD as well as dynamic adjusting of asset portfolio volatility.

I. Model
I assume that asset values follow a jump diffusion process with an exogenous interest rate.\(^7\)
\[
A_{j,t} = A_{j,t-1} e^{X_{j,t}} \tag{1.}
\]
\[
X_{j,t} = (\mu_X \Delta t + \mu_J Y_t) + \sqrt{\sigma_X^2 \Delta t + \sigma_J^2 \varepsilon_{j,t}} \tag{2.}
\]
where \(\varepsilon_{j,t} \sim N(0,1)\). Eq. 1 specifies firm j’s asset process as a stochastic process governed by a
jump-diffusion process \(X\). Let \(\sigma_X\) represent correlated asset volatility, and \(\mu_X, \mu_J\) represent
mean of the firm’s asset process and jump process respectively. \(\sigma_J\) denotes the volatility of the
jump process. \(Y_t\) represents jumps.
A firm will default as soon as its assets become unable to cover its liabilities due. We specify
firms’ threshold of default as a sum of long term liabilities, discounted to their present value, and
short-term. Let short term ToD equal to short-term liabilities \(D_{j,t}^S\).
\[
K_{j,t}^S = D_{j,t}^S \tag{3.}
\]
Long-term liabilities equal the discounted future amount due at the rate of the firm’s cost of
capital r.
\[
D_{j,t}^L = e^{-r(T-t)} D_{j,T}^L \tag{4.}
\]
Firm’s assets are allowed to drop below its liabilities prior to the due date. Specifying a flexible
ToD enables us set a benchmark below which asset values should not be allowed to drop at any
time. Defining a credit accessibility parameter S enables us to inject such flexibility into a firm’s
ToD.
\[
S_{j,t} = e^{-(\beta_1 Z_j)} \tag{5.}
\]
Credit availability spread S is a function of individual firm’s Beta and a system-wide factor Z. Z is
a function of the volatility index VIX, change in the yield curve ratio, and a yield curve ratio
dummy variable reflecting yield curve inversion or a lack of thereof. The yield curve ratio is
specified as a ratio of the short end of the yield curve to the long end of the yield curve. If the
ratio shows values greater than one, the yield curve has inverted.
\[
Z_t = \sqrt{(VIX_t + \Delta YCR_t^2)} + YCRD_t \tag{6.}
\]
As \(\lim_{Z \to 0} S = 1\) and \(\lim_{Z \to \infty} S = 0\), and thus, S approaching one signals that assets are allowed to drop in
value by a hundred percent. S equal to zero brings ToD to the present value of long-term
liabilities, so that \(K_{j,t}^L = e^{-r(T-t)} D_{j,T}^L\).

\(^7\) Contact author for details.
Let the long-term threshold of default be a function of credit availability spread, cost of capital, long-term debt obligations, and time remaining to maturity.

\[ K_{j,t}^L = e^{-\left(\delta_{j,t} + \gamma \right) T_{j,t}} D_{j,t} \]  

(7.)

Finally, I define the threshold of default as a sum of the short-term and the long-term thresholds.

\[ K_{j,t} = K_{j,t}^S + K_{j,t}^L \]  

(8.)

I use the flexible threshold of default to target a company’s optimal distance to default. I allow for fluctuations within a set of boundaries, effectively minimizing default probability to values nearing zero.

Let \( A_{j,t} \) denote a firm’s asset value, and \( K_{j,t} \) denote the threshold of default. Let \( \sigma_A \) represent correlated asset volatility. Then firm j’s distance-to-default at any point in time can be specified as:

\[ DtD_{j,t} = \frac{A_{j,t} - K_{j,t}}{\sigma_A} \]  

(9.)

In order to achieve a steady DtD, I have to manipulate the firm’s asset portfolio to reflect a specific credit risk strategy - either curbing or expanding portfolio asset volatility parameter. In fact, in this study I allow DtD to fluctuate in-between pre-specified interval boundaries, always reverting back to the target DtD.

Suppose that ToD contracts/expands, and concurrently asset volatility doesn’t change, leading to an increase/decrease in credit risk. The following method attempts to address such instabilities. In particular, manipulating asset portfolio volatility enables me to target and maintain an optimal DtD.

Firstly, higher market volatility, leads to higher levels of the VIX index, as well as higher asset value volatility correlations. As a result, ToD contracts, making the gap between asset value and ToD smaller. High levels of the YC ratio or an YC inversion have a similar effect. If at the same time we manipulate asset portfolio such that its volatility contracts, we’ll be left with a steady distance-to-default. Conversely if ToD expands, making the gap between asset value and ToD larger and we extent asset portfolio volatility, we again, achieve a desirable DtD.

Let the target DtD be set at two standard deviations from the ToD. Also, let the trigger points ( \( b_L, b_U \) ) be set at one and three standard deviations from the ToD. If the DtD becomes smaller than one, a trigger will be sent to signal asset volatility decrease such that we achieve the target DtD. If the DtD becomes larger than three, we can increase asset riskiness further to capitalize on a favorable credit environment, adjusting DtD back to its target.

\[ b_L \leq DtD_{j,t} \leq b_U \]  

(10.)

Adjusting portfolio asset volatility requires taking into account individual asset volatilities, correlations between individual assets’ volatilities as well as proportions of individual assets in the overall asset portfolio. In order to manipulate the asset portfolio volatility I can either change the portfolio composition or readjust weighting of the existing portfolio constituents. I calculate the asset portfolio volatility as a correlated weighted average of the individual assets’ volatilities.

\[ \sigma_A = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \rho_{i,j} \sigma_i \sigma_j \]  

(11.)
Supposing that individual asset volatilities are exogenous, as is their correlation matrix, we shall manipulate asset weights so that we achieve desired portfolio asset volatility. Such volatility is determined by triggering readjustment of the asset portfolio so as to satisfy the target DtD.

Expected pair-wise correlation matrix $\rho_{ij}$ changes with changes in system volatility, Das (2006). Specifically, an increase in market volatility increases correlations between asset volatilities. Let us denote current correlation coefficient matrix as $\rho_{0,ij}$. Conditional on $-1 \leq \rho_i(1 + E[D_{t+1}]) \leq 1$, I determine forward-looking correlations as a function of the current correlation matrix and a delta of the volatility index.

$$\Delta_{t,i} = \tanh(\gamma_i \Delta VIX_i)$$

Individual assets in the asset portfolio may react differently to the changing market conditions. Hence I include a parameter $\gamma_i$ as a sensitivity-reflecting measure. For example, cash will be virtually insensitive and its correlation coefficient will experience no or very minimal change as a response to the current market conditions. I let $\gamma_i \in [0,1]$. Expected change in the correlation coefficient can be calculated using a probability density function $f(\Delta VIX_i)$ reflecting the market risk conditions.

$$E[\Delta_{t,i}] = \int_{-\infty}^{\infty} \Delta_{t,i} f(\Delta VIX_i) d\Delta VIX$$

II. Data, Results

Data sets used in this study run from January 1990 to October 2010, totaling 5193 observations each. Credit availability spread, specified above, has been calibrated with a set of daily observations of: (1) the VIX index, (2) the 3-month interest rate representing the short end of the yield curve, and (3) the 10-year interest rate as a proxy for the long end of the yield curve.

Using the yield curve data, we shall construct a yield curve ratio (YCR) that reflects the short end relative to the long end. Ratio values higher than one represent yield curve inversion. Further, we model the credit availability spread taking the first difference of YCR data. Once the yield curve inverts, we add this effect to the model in the form of a dummy variable, showing one for inverted yield curve and zero for non-inversion.

The following figure shows the basic data set of variables used in this paper, ie. the VIX index, changes in the yield curve ratio, and yield curve inversions. Individual firm’s credit availability spread $S$ is a function of system-wide conditions. Those are reflected by parameter $Z$ as defined above. In periods of yield curve inversions and extreme market volatility the chart shows values greater than one. Suppose a portfolio consists of three firms: A, B, and C. We assume they are alike in all respects but the riskiness of returns. Firm A has a beta equal to one, firm B has a beta of two, and firm C has a beta of 0.5. The following figure shows the effect of varying betas on the credit availability spread $S$.

Over time we observe variations in spread $S$. Along our data set, we can identify a few periods of credit distress. Each of the three firms reacts to the systemic disruptions at identical points in time. Specifically, $S$ dips for all firms in times of the yield curve inversion and/or increased market volatility. We can distinguish individual magnitudes of reactiveness dependent on firms’ riskiness. Recall that $S$ can take on any value from and interval between zero and one. The least risky firm (C) shows values of $S$ close to one at most times. The most risky firm B, on the other hand mostly shows $S$ levels at around 0.4.
At the outset of the asset process simulation exercise asset values had been set uniformly at one hundred creating a fairly large distance to default for all three firms. At the same time we could observe fairly low market volatility in the early nineties. Both of these effects combined result in consistently increased levels of the volatility parameter peaking at around 40%.

This paper sets out to formulate a method for targeting and maintaining a desired distance-to-default. Once a firm’s asset portfolio value moves considerably close or far away from its threshold of default (as defined above), we receive a signal of an asset volatility adjustment requirement. The asset volatility parameter adjusts to keep the DtD moving within acceptable target boundaries. Figure three shows asset volatility parameter adjustments.

The charts on the left hand side of figure three show asset value standard deviation sigma generated by a random number generator given firm risk-specific means and standard deviations. I used a random normal value generator with mean values reflecting A, B, and C’s riskiness (0.1, 0.18, and 0.05 respectively). The right hand side shows asset volatilities for firms A, B, and C after they have been adjusted applying the target DtD method specified above.

Firstly, let us consider the charts in the first row of the chart matrix. It shows asset volatility parameter for firm A. Firm A is assumed to have a beta equal to one. Comparing the first row charts reveals that asset portfolio volatility adjustment would have been made downwards in 2007/2008 and 2009, corresponding to an end of an asset bubble period.

Let us now consider firm B whose riskiness, measured by beta, had been set at two. We see...
more triggers of asset adjustments than was the case with firm A. There are more plentiful downward volatility adjustments for B. All adjustment time periods occur at the end of asset bubble bursts. The asset volatility parameter adjustments are noticeably more pronounced. During the last decade, adjusted sigma parameters are frequently reaching zero and subsequently, in times of low systemic stress, expanding to and beyond their original position.

**Figure 3: Asset volatility parameter before and after adjustment for a steady DtD for A, B, C.**

Finally the lowest risk firm C exhibits a volatility increase starting in 2008, persisting until the end of our data set. This result suggests that C’s limited reactiveness to market conditions held up its asset value processes in part independent of the market conditions and thus its distance to default signaled values favorable to increasing asset portfolio volatility.

Asset process simulation method used in this paper is governed by assumptions about the asset value volatility parameter. As the asset value processes are modeled to include jumps, the jump volatility parameter was set at a mean of 0.1 with a time-varying random normal component. The jump intensity parameter lambda was set at 0.75, and the jump frequency was generated as a random variable drawn from the Poisson distribution. Interest rate used is exogenous, set at 0.05.

The following figure shows a randomly chosen simulated asset value processes for firms A (blue), B (green), and C (red). Asset value processes in the original form reflect the original random volatility parameters. Thus the first chart, in the upper part, has asset processes displaying relatively low volatility up until the time of the first yield curve inversion. The processes fan out afterwards.
Once we have applied the DtD signaling method that leads to adjustment of asset portfolio volatility, we observe greater asset volatility at times of smaller systemic disturbances. Looking at the lower chart, it is represented roughly by the first ten years in our data set. At times of higher market volatility and yield curve inversions, asset processes reflect the new adjusted volatilities and their paths become relatively more suppressed.

Lastly we plot firms’ asset processes together with their respective thresholds of default. In figure five we assume that the firms’ liabilities become due at the end of our dataset. The face value of the firms’ liabilities at maturity is arbitrarily set at one hundred. Note that this chart contains both – an elastic ToD as introduced in the model above, as well as a rigid default barrier given by the present value of the firm’s liabilities. We can see firms’ distances to default at different times-to-maturity.

Top chart of figure five shows the original asset values with their corresponding elastic as well as an inelastic ToDs. The bottom chart shows the volatility-adjusted asset processes. Note that the ToDs are identical in both charts of figure five. Note that in this particular simulation scenario one of the three firms’ defaults. Here, the asset value process of the most risky firm B drops below its ToD. This happens under the original assumptions about asset value volatility as well as with the new adjusted volatilities. However, there is a difference. In the original scenario with no asset volatility adjustments, firm B’s asset values fall beneath its elastic ToD already in 2004 and stay there until the maturity time. Volatility adjustments allow firm B to preserve asset values above its ToD longer. B’s asset process, in fact, displays a tendency to reverse the drop back above the default threshold in time before the firm’s liabilities become due. Even though firm B was unable to avoid default, a volatility-sensitive asset management approach would have increased its chances of survival.
III. Conclusion

This paper proposes a credit-risk management method for individual firms. The model put forward reflects existing market conditions. We can utilize it to determine a firm’s threshold of default reflecting market volatility and credit accessibility conditions. The dynamics of the default threshold allow it to expand with low (and contract with high) system volatility and high (low) credit accessibility. I use the threshold of default to measure firms’ existing distance to default. Distance to default is central to credit risk management as it specifies the firm’s asset value relative to the threshold. Asset values fluctuate continuously and so does the above-specified threshold, both reflecting current market conditions.

This study uses macroeconomic and market data spanning over the last twenty years. I applied it to the DtD targeting method assuming that a hypothetical firm’s liabilities mature at the end of this data set. Calibrating the model with the real-world data resulted in episodic signals for asset portfolio volatility adjustments. Running Monte Carlo simulations on mock asset portfolios with and without adjustments resulted in lower frequency of asset values crossing the ToD for the adjusted portfolios. The volatility-adjusted portfolios compared to the unadjusted ones achieved the desired effect of maintaining the firm’s DtD in its pre-specified boundaries, effectively decreasing firms’ default probability.


Crosby, Peter, and Jeffrey Bohn. 2003. Modeling Default Risk. KMV Corporation


Superannuation Fund Communications in Good and Bad Times

Malcolm Smith*
School of Accounting, Finance and Economics
Edith Cowan University
270 Joondalup Drive, Joondalup  WA  6027
malcolm.smith@ecu.edu.au

Paul Gerrans
University of Western Australia

Jacqui Whale
Edith Cowan University, Western Australia

*corresponding author
Abstract
We examine the content and presentation of the first statement to members appearing in a Superannuation Fund’s annual report (typically made by the Chairperson or the CEO). We reflect on the words and themes employed, and the clarity and readability of communication. Our focus is on the respective differences between statements made in the year prior to the global financial crisis (GFC) and those made during the GFC.

A content analysis of the CEO statement was conducted for 81 annual reports in each of the two years - respectively ‘good news’ and ‘bad news’ years.

The readability of statements (as measured by the FLESCH reading ease index) is surprisingly consistent. Readability scores for ‘bad news’ statements do not differ significantly from those for ‘good news’. Although the themes appearing in the statements remain relatively constant, their ordering and the emphasis accorded them differs greatly. ‘Good news’ statements focus on fund investment performance, while ‘bad news’ statements are oriented more to ‘market performance’. This is consistent with the literature which suggests that poor performance will be attributed to external factors. Compared with company annual reports, the superannuation fund statements tend to adopt a more ‘folksy’ style, reflecting both the different audience and purpose of the report.

If message consistency is an attribute of successful communication, then there is room for improvement in the writing of the statements. The role of the market performance and its impact on fund investment performance needs highlighting in good and bad return periods. Similarly the emphasis on investment time horizons, particularly descriptions of superannuation as a long-term investment, also needs reinforcing in good and bad times. This recommendation is consistent with ASIC’s move in 2008 to ensure funds include medium to longer term performance data in annual reports and/or member periodic statements.

The paper contributes to the literature on readability in corporate communications by examining the annual reports of Superannuation Funds experiencing a ‘bad news’ year due to the GFC, and comparing them with the previous years’ ‘good news’ reports.

Keywords Superannuation Fund Report, Good news/Bad news, Themes.
was in its infancy with 40 percent employee coverage (APRA, 2007), and industry assets of $41 billion (Neilson and Harris, 2008). By 2008, assets totaled $1.1 trillion (APRA, 2008) with 91 percent of those employed with superannuation coverage, and an average balance of $70,670 (ABS, 2009).

### Table 1: Returns in 2007 and 2008 of flagship investment options

<table>
<thead>
<tr>
<th>Fund type</th>
<th>2007 Return (%)</th>
<th>2008 Return (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Max</td>
</tr>
<tr>
<td>Corporate Funds (n=8)</td>
<td>16.3</td>
<td>18.2</td>
</tr>
<tr>
<td>Industry Funds (n=42)</td>
<td>16.4</td>
<td>21.0</td>
</tr>
<tr>
<td>Public Sector funds (n=13)</td>
<td>15.3</td>
<td>17.7</td>
</tr>
<tr>
<td>Retail Funds (n=14)</td>
<td>15.3</td>
<td>25.9</td>
</tr>
<tr>
<td>Total (n=77)</td>
<td>16.0</td>
<td>25.9</td>
</tr>
</tbody>
</table>

* A return for both years could not be identified for four funds in the sample

This paper is concerned with the information and messages that retirement savings fund members received during the GFC, using Australian superannuation funds as a case study. At the broad level this represents a potentially extensive topic given the myriad communications members receive. These include direct print materials from the fund, the fund’s online website, and also the various commentaries in the media (television, magazines, newspapers, and online). The communication that registers most directly, and also mandatorily, is arguably the annual report which, in the majority of cases, accompanies the member’s annual balance statement. This paper focuses on the annual report and more specifically either the Chairperson’s or CEO’s statement which introduces the report. The analysis examines the words, language, and themes in these statements contained in the June 2006/07 and June 2007/08 annual reports.

The specific research questions to be investigated in this paper are:

- How have superannuation funds used the Chairperson’s statement to communicate bad news and good news to their members?
- Is there a consistent structure to the Chairperson’s statement?
- How readable are the statements?
- What language and key themes are evident in statements and are they consistent?
- Is there evidence that message style, construction, and consistency, is different in periods where superannuation funds experienced good (2007) versus bad (2008) market conditions?

### I.1 Requirement for Periodic Statements and Annual Reports

The Corporations Act (S1017D) establishes the requirement for the issuer of a superannuation product to provide at least one periodic statement to fund members. The specific information required includes: opening and closing balances; transactions data; return on investment; termination value; and details of any changes in the circumstances affecting the investments. Beal (2008, p.50) notes that statements often include a range of additional general (investment strategy returns, insurance, and complaints mechanisms) and specific (insurance benefits, beneficiary details) information “in a fairly logical sequence” (Beal, 2008, p.50). Like periodic member statements, annual reports also appear to follow a “logical sequence” with a key feature the importance given to the statement from the Chairperson, CEO, or combination of the two.

The Corporations Act (S1017DA) establishes requirements for information in annual reports “relating to the management, financial condition and investment performance of the entity and/or
of any relevant sub-plan”. The Corporations Regulations (Subdivision 5.6 of 7.9) provides further information requirements.

I.2 Are Annual Statements and Reports Important for Members?

Several surveys have questioned how members utilize their fund’s annual report and suggest somewhat contradictory results. The Association of Superannuation Funds Annual Superannuation Survey (Balogh, 2008) suggests 79 percent of respondents considered their superannuation fund communicated with them well and 86 percent could mostly or easily understand their transaction statements. The same survey found that 46 percent of respondents did not know the actual percentage return of their fund for the 2007/08 return. The Investment Trends (2007) survey also presents results suggesting satisfaction with the information provided by funds in superannuation statements. The sample focused on those who had switched superannuation funds in the previous 12 months. In separate assessments on information relating to performance, fees, risks, and the market the highest proportion of respondents reporting they considered the information poor, very poor or not used was 23 percent. When asked what could improve the statements, ten percent indicated more information, 11 percent indicated statements that were clearer or in plain and simple English, and 14 percent indicated nothing or that they were fine.

The Australian Securities and Investments Commission (ASIC) suggest periodic statements were “probably the most important bit of disclosure for ongoing members of a super fund” (Cooper, 2007, p. 4). The Investment Trends (2007) survey also confirms that the majority of members (65 percent) read the periodic statements and 67 percent indicated that they understood them well or very well.

I.3 Who is the Audience of a Superannuation Fund Report?

Who actually is the intended audience of the annual report is an important question as it goes to the heart of the issue of how members of funds are viewed and how the statements can be judged. Donald (2008) examines the different way funds and members are treated within the superannuation regulatory scheme which is of direct relevance to the question of the purpose of the annual report. At one extreme is the view of members as the “hapless and vulnerable beneficiary of Equity lore” and at the other is the “competent economic actor” (Donald, 2008, p.5). In the former traditional model, regulation seeks to ensure trustees administer as per the paramount trust instrument. Donald suggests that although some view this as anachronistic in the light of what modern superannuation funds have become, both the courts and system reviews have failed to overturn the trust law basis for superannuation funds as is applicable to general trusts.

The requirements placed on superannuation trustees from trust law have, however, been modified by employment and contract law. This reflects that for the majority of members, involvement in a superannuation fund is not voluntarily, as in the traditional trust, but rather a consequence of their employment and the Superannuation Guarantee legislation. Donald (2008) suggests however that the influence of this employment contractual context should not be overstated as many members still have the “‘vulnerability’ that attracts obligations in other fields” (Donald, 2008, p.7).

Given the compelling size of the superannuation industry a third “public model” is suggested by Donald (2008). Given the inter-relationship between pension entitlement and superannuation income as sources of retirement income, the public interest case can be mounted. Acknowledging this, the suggestion is that superannuation funds and trustees should also be subject to the principles governing public officials including administrative law. Whether this is
appropriate or not is not the focus of this paper, though it may be regarded as the least compelling model. However, it is within this framework that Donald (2008) suggests the requirements of the Corporations Act of certain prescribed information can be viewed. Importantly, “much discretion is left in the hands of the trustee” (Donald, 2008, p.12). This is not to suggest that the three models are mutually exclusive. The Australian regulatory system incorporates aspects of each model. From the member perspective there is evidence both across fund categories and within a fund that they are populated by both the traditional model’s beneficiaries and investors. The proportion of assets in the default investment option within each fund type provides an imperfect measure for each. Retail funds have the lowest proportion (23 percent) and industry funds the highest (74 percent) in default options (APRA, 2009) suggesting they are populated by a greater proportion of investors and members respectively. Industry funds, notwithstanding the increasing number attaining public offer status, remain populated with compulsory (beneficiary) members whereas retail funds could be reasonably characterized with more voluntary (investor) members. It should be noted that this need not infer anything about member sophistication.

Confounding the investor characterization, however, is the low proportion of members who exercise both investment choice and choice of fund. Sherry (2008) suggests that 80 percent of fund members do not exercise choice. This is comparable to overseas evidence. For example in the U.S., Mitchell, Mottola, Utkus and Yamaguchi (2006) show the majority of defined contribution plan participants do not trade and remain in the plan’s default investment option. It is here where the different constituents of superannuation funds, which may well align with the different categories of funds, may determine or influence the focus of member communications. This may be true generally and for the annual report, including the Chairperson’s or CEO’s statement, specifically. For example, a member by choice of a retail Master trust can be contrasted with the compulsory members of Industry or Corporate funds receiving the minimum contributions, who in the majority remain in the default option. The former, voluntary members of retail products, or members by choice of public offer industry funds for that matter, accord with the investor model. The latter, compulsory and largely passive member groups accord with the beneficiary model.

Putting aside the “public model”, it is worthwhile to consider the different information sets that may be considered as necessary given the “member” and “investor” models. What assumptions can be made about member information requirements, knowledge levels, and to what extent are these reflected in the Chairperson’s and CEO’s message?

II. Review of the Literature

Readability definitions incorporate a number of dimensions. These dimensions include the ease in understanding a text, the degree to which the text is compelling and comprehensible, and the success of the reader in understanding the text (DuBay, 2004). Though the history of assessing readability is long, a single computational model for measuring all aspects of readability is not available (Pitler and Nenkova, 2008). Instead, most assessments choose a single factor of relevance for a particular audience (Pitler and Nenkova, 2008). This paper belongs to this latter approach with a focus on assessments of readability in corporate communications.

In an ideal world corporate communications will reflect financial reality. Unfortunately the use of jargon is commonplace, and goes well beyond that which is necessary for technical accuracy. Whether this matters or not will depend on the way in which the information is presented, because the means of conveyance will tell us something about the communicators and the well-being of their organizations.
Smith & Taffler (1992a, 2000) suggest that the readability of accounting narratives is linked to financial performance, and that the use of key standard words and phrases is predictive of poor performance; deteriorating financial conditions are often accompanied by increasingly opaque communications. Clear and transparent communications reflect well on both the communicators and the organizations that they front. Trust and respect can be fostered with such a device. On the contrary, if miscommunication is interpreted as a deliberate attempt to mislead, then that trust may be lost. The absence of clear messages may signal that there is something seriously wrong with both the management of the communication process and the management of performance. Effective communication demands that the recipient interprets the message in the same way as the sender. The more trustworthy the sender then the more credible and reliable is the associated message. However, the communication process can break down if the reader cannot understand or misinterprets the content, or because the sender has deliberately constructed a misleading message. The message might contain misleading content (downright misinformation), or might be presented in a way which makes it more difficult to understand through poor readability and/or the overuse of jargon. Smith & Taffler (2000) show that ‘bad news’ is normally associated with references to ‘losses’, ‘contraction’, and ‘resignation’ while ‘good news’ is associated with talk of ‘profits’ and ‘growth’.

The complexity of financial communications increases the potential for impenetrable narratives, full of technical terms, and if the user is not able to read, understand and interpret the content, then opportunities for communication breakdown will arise. Smith & Taffler (1992a) suggest that firms will want to signal their superiority to the market, both through excellence of performance and through transparent disclosures. But there may be differences in approach depending on whether ‘good news’ or ‘bad news’ is being imparted; if there is no ‘good news’ to be disclosed, then some firms might try to obscure the fact by conveying messages which are either unclear or incorrect. The implication is that organizations will signal their good performance while trying to obscure poor performance; this might be combined with attributions which suggest that ‘good performance’ is all down to brilliant management, while ‘poor performance’ can be blamed on economic and market conditions. Poor performance is normally accompanied by a gloomy, pessimistic narrative, while excellent financial numbers are complemented by self-congratulatory and optimistic words. Beattie, McInnes, and Fearnley (2004) provide a review of the analyses of annual report narratives and propose a metric to measure not only the quantity but also the quality of financial disclosures in annual reports.

The GFC presents a challenge not only in its direct impact on member accumulated savings, but also the retirement income planning system as a means to deliver member’s desired standard of living in retirement. The existing literature focuses on firms’ corporate performance rather than superannuation funds. This Analysis is therefore timely given that membership of such funds is compulsory for workers in Australia. Recent research from de Grip, Lindeboom, and Montizaan (2009) has demonstrated the negative impact of changes to pension system entitlements on the mental health of workers in the Netherlands. A parallel can be drawn with the GFC which has been an external shock to worker retirement entitlements which dependent upon member characteristics, age or time to retirement in particular, will have direct impacts. Thus an analysis of fund messages provided to members in the wake of the GFC, is opportune and will provide superannuation funds with an industry wide analysis for reflection.

II.1 Differences in Annual Report Purpose

The previous research into company annual reports is useful but it also highlights a major point of difference in the audience of the reports. For corporations, the reports are unambiguously for
current or potential investors. They have been described as “platforms for preaching philosophies and touting themselves and their companies” (Ingram and Frazier, 1983, p.49). This highlights an inherent tension in the use of the annual report as a marketing tool rather than predominantly a disclosure or educative instrument. Given a fund membership dominated by “beneficiaries”, it is arguably more beneficial to see the annual report role as primarily educative.

Australians have a comparatively high rate of share ownership with an estimated 38 percent owning shares directly. However, superannuation is held by a much larger proportion, 79 percent of households (Reserve Bank of Australia, 2009). Further, average household superannuation wealth, at 15 percent of total assets, is more than double the seven percent that direct equity comprises.

Direct equity ownership serves a variety of investment purposes for individual investors. Australian Stock Exchange (2006) identifies that 33 percent of investors invest in equity “to accumulate wealth”, 12 percent for “diversification” and 15 percent of individual direct equity investors invest in equity for retirement. There is therefore an obvious overlap of superannuation members who are also direct equity investors, though the larger proportion of superannuation members do not have this experience. Even for those holding equity, 42 percent hold three or fewer companies in their portfolio.

There are other similarities and subtle differences between the two audiences. Both the fund and corporation aim to provide a return for members and owners respectively. As noted above, an equity investment may be for a variety of objectives and investors can access the return immediately. In contrast, a fund member’s objective is retirement savings and importantly they can only access when they reach their preservation age, which for the largest proportion of workers is 60.

II.2 Readability and Understandability

The survey evidence suggests two major areas of concern in the manner of the disclosure of superannuation information: clarity and understandability. This paper addresses issues of clarity in terms of the difficulty of text employed and the reading ease of the narrative.

Early research (e.g., Smith & Smith, 1971; Healy, 1977) focused on the comprehensibility of footnotes in annual reports, raising serious doubts about both their complexity and readability. Related research has sought to develop alternative measures of readability, based on the difficulty of the narrative, the length of the sentences and the length of the words employed. Smith and Taffler (1992b) note that all the measures combine two basic features:

- Word Length (W) – related to the speed of recognition and acting as a proxy for the difficulty and/or obscurity of the language;
- Sentence Length (S) – related to the recall of words in immediate memory, so that meaning and context allow the message to be conveyed.

Alternative measures arise because of the different possible ways of measuring average word length and average sentence length. The LIX index distinguishes between words which are greater than or smaller than six letters in length. The FOG index (Kwolek, 1973) focuses on words deemed to be ‘hard’ from a prescribed dictionary. The FLESCH index uses average number of syllables. Of the available measures the LIX measure (Anderson, 1983) is the easiest and most reliable to compute, but the FLESCH index (e.g., Still, 1972) remains the most popular in the literature, hence the use of the latter here.

The FLESCH index defines W (word length) as the number of syllables per 100 words, and S (sentence length) as the total number of words divided by the total number of sentences. The FLESCH readability formula is then computed as: FLESCH = 206.835 – 0.846W – 1.015S. The
calculation thus represents a deduction from the base constant for both word and sentence complexity, so that the higher the score the easier the readability. A FLESCH score of less than 30 would signify a difficult technical narrative. Smith and Taffler (1992b) report FLESCH scores in the range 24.8 (Bad) to 49.3 (Good) for their sample of Chairperson’s narratives, with a mean score of 34.2.

The LIX measure is calculated as LIX = W + S, where W is the percentage of words of seven or more letters, and S is average number of words per sentence. The lower the score the easier the readability, so that LIX scores in excess of 60 would signify difficult technical narratives. The resulting indices provide reliable reflections of the difficulty of the text, but ones that still have limitations, since word length and sentence length are both used as proxies for difficulty. For example, efforts to improve the understandability of the text, through the addition of subordinate explanatory clauses, would actually reduce the readability because the resulting sentences would be longer. Also, and importantly, the readability measure is a single score which makes no reference to the reader, nor to differing levels of experience, maturity or sophistication among the readership.

The development of a measure which reflects variations in the abilities of the readership requires a different approach. It must be interactive, and necessarily more time-consuming to construct and conduct. The CLOZE procedure (e.g., Taylor, 1953) is one such option which has been well researched. The CLOZE method seeks to predict the understandability of a complete narrative based on comprehension of parts of the whole. Subjects are provided with a mutilated text and asked to predict the missing components based on the surrounding context. In practice versions of the text are prepared with every nth word omitted (usually every fifth word or every tenth word). The easier the text is to comprehend the easier it will be to fill in the gaps, so that the percentage of correct insertions can easily be computed. Familiarity with the nature of the texts should allow subjects to record higher CLOZE scores, but groups of similar sophistication would be expected to record similar scores. Overall scores are generally low, since only exact insertions are counted as correct.

Limitations still exist; for example, the guessing of an exact response may not necessarily correspond with a complete understanding of the message. More recent alternatives in the area, notably the Sentence Verification Test (SVT) developed by Royer (2001), have sought to develop specific questions on the content of the narrative to determine the real level of understanding.

III. Sample Overview and Coding Methodology

A sample of funds covering each of the main superannuation fund types was constructed. While the peak of equity market declines occurred in March 2009 the impact of the GFC on fund returns was evident in June 2008 with the largest ever reported annual negative returns. Hence the 2008 report was taken as a source of bad news and the 2007 report a source of good news. The sample breakdown is presented in Table 2. To be eligible for analysis the fund report must have been reporting for the June 30 financial year. The Chairperson or Chairperson/CEO statements were extracted for 2006/07 and 2007/08 from the electronic versions available from the fund website. In a small number of cases the fund was contacted using the “help centre” email address. In 26 cases a selected fund was dropped from the analysis because a report for both years was not available. The final sample of 162 statements, two for each of the 81 funds, was loaded into NVivo for text coding. A full list of the funds included in the sample is provided in the Appendix.
Table 2 Sample Overview

<table>
<thead>
<tr>
<th>Fund Classification</th>
<th>n</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corporate</td>
<td>8</td>
<td>3131</td>
<td>139</td>
<td>10059</td>
</tr>
<tr>
<td>Industry</td>
<td>42</td>
<td>4197</td>
<td>139</td>
<td>28834</td>
</tr>
<tr>
<td>Public Sector</td>
<td>13</td>
<td>7114</td>
<td>1041</td>
<td>24539</td>
</tr>
<tr>
<td>Retail(^1)</td>
<td>18</td>
<td>3750</td>
<td>47</td>
<td>23900</td>
</tr>
<tr>
<td>Total</td>
<td>81</td>
<td>4469</td>
<td>47</td>
<td>28833</td>
</tr>
</tbody>
</table>

Statement Completed by

<table>
<thead>
<tr>
<th></th>
<th>2006/07</th>
<th>2007/08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chairperson</td>
<td>51</td>
<td>46</td>
</tr>
<tr>
<td>Chairperson &amp; CEO</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>CEO</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Managing Director/Directors</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>President/Trustee/Board</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>81</td>
<td>81</td>
</tr>
</tbody>
</table>

\(^1\) Asset value not established for one fund

III.1 Coding

The statements were coded to themes at the paragraph level. Rather than select the major theme for a particular paragraph, each paragraph was permitted multiple themes. The paragraph number was also coded, allowing an examination of theme sequencing. Two coders were utilized for the 162 statements.

The first round of coding involved reading the full statement from a sample of twelve funds. This helped establish the major themes and develop preliminary rules of classification. In the second round of coding each coder independently coded a share of the total sample. When all statements were coded, theme definitions were collectively refined with rules of theme classification revised. In the third round of coding each theme was checked to the established rules to ensure that each paragraph was appropriately classified.

III.2 Themes

The coding process identified four broad categories of themes and 16 more refined themes.

III.2.1 External References

The first category of themes refers to discussion of factors external to the fund itself. These themes refer to the broader environment in which a superannuation fund exists. Discussion of the performance of various financial markets was a major theme in this category. The **Market Performance** theme includes discussion of share, bond and property markets, either domestic or global. A common example in 2007 is:

“Domestic equity markets enjoyed a stellar year in 2006/2007, with the All Ordinaries Index gaining more than 25%” (The Bendigo Superannuation Plan, 2007).

In 2008 this had changed to:

“These results were mainly caused by negative returns from overseas share markets (-21%) and the Australian share market (-14%)” (AUST (Q), 2008).

In 2008 there was also reference to other markets or general economic factors related to the global financial crisis, such as:

“High oil prices and the US housing collapse have caused a flow-on effect in financial markets throughout the world” and “Inflationary concerns arose after world oil prices exceeded US$140 a barrel for the first time” (The Bendigo Superannuation Plan, 2008).
This theme specifically excludes general information or advice about share markets such as:
‘equity markets generally outperform in the long run’. Such statements were coded to an ‘Advice & Education’ theme.

Many statements referred to the superannuation regulatory framework. The **Superannuation Environment** theme included references to legislative changes, notably ‘Better Super’ in the 2007 report, and references to the legislation governing superannuation. Common examples are:

> “An important goal for the Trustee this year has been adapting the Fund's super products to ensure compliance with the Federal Government's Better Super System” (Accountants Super Fund, 2007), and
> “The outcome of the recent review by the regulator, the Australian Prudential Regulation Authority (APRA), was very pleasing and members can be satisfied” (Catholic Superannuation Fund, 2008).

**III.2.2 Fund**

The second category of themes covers messages about the fund itself, including: how the fund is operated; how its investment options performed; the goals of the fund; discussion of the fund’s profile; and the products and services of the fund. This is the largest category of themes.

The **Board** theme includes references to the composition of the Board of Directors; Board and senior staff changes, including new appointments; Board meetings; and the responsibilities of Trustee Directors. For example:

> “As part of our Board’s continuing efforts to improve our Fund, we made the decision during the year to create the position of Chief Executive Officer” (Club Plus Superannuation Scheme, 2007), and
> “As at 30 June 2007, there were three new faces on the Trustee board” (IAG & NRMA Superannuation Plan, 2007).

The **External Endorsement** theme categorizes statements where the name of an external and independent body was used to show the Fund’s performance and/or products in a good light. Most often this was in relation to ratings and awards, such as:

> “We were also a Finalist in Super Ratings’ Fund of the Year awards in November 2006 and received The Heron Partnership’s Top Rated Retail Product award” (AGEST, 2007), and
> “Our investment performance and the quality of our funds continues to be recognized publicly with BT being awarded Fund Manager of the Year at the 2006 Australian Financial Review Smart Investor Blue Ribbon Awards” (BT Lifetime Super, 2007).

Statements referring to fee levels or strategies focused on reducing cost levels were coded to the **Fees** theme. This includes specific reference to fees (low fees) or to ‘low cost’ and economies of scale. For example:

> “This reflects our commitment to excellence and low cost for members across all our products, both before and after retirement” (AGEST, 2007), and
> “Equally significantly, we’ve been able to implement these features while maintaining low fees. And its fees, coupled with investment returns that make the most difference to account balances” (AustralianSuper, 2007).

The **Investment Performance** theme complements ‘Market Performance’. Here, the fund’s investment option returns are specifically referred to, either in general terms, for example:

> “On a positive note we have seen good returns over the past 12 months and we can thank our independent Investment Committee for keeping our Investment Managers performance to the maximum benefit of members” (Astarra Superannuation Plan, 2007),

or more specifically:
“I am pleased to confirm that investment returns for members have been excellent across all investment options with the default option, Managed Growth (70/30), achieving a return of 15.4%, p.a. to the year 30 June 2007” (Accountants Super Fund, 2007).

Discussion of the fund’s overall investment strategy or the asset allocation of specific investment options was coded to the Investment Strategy theme. For example:

“On the subject of investments, the Board recently approved small changes to the target investment mix of three of our blended options - High Growth, Medium Growth and Stable Growth” (Asset Super, 2007).

In 2008, references to Investment Strategy could often be characterized as ‘holding the line’ on strategy. For example:

“I note that the Plan was (and remains) invested in appropriately diversified investment strategies. This includes a significant exposure to high quality, liquid assets, which help to minimize the risks associated with any particular investment strategy or investment manager.” (Qantas Superannuation Plan, 2008).

The theme also included references to the outsourcing of investment management; investment advisors; strategic asset allocations; and the monitoring and review of investments.

The broadest theme in the Fund category was Operations, Goals & Objectives, including references to: the administration of the fund; member communications and research; staffing and management structure; overall operational strategy; fund focus, objectives and aims; and evaluation of products and services. A general example which addresses fund objectives is:

“Looking forward, we aim to be recognized as the “fund of choice” for those public sector employees who have choice in their workplace and to continue to seek default status from employers in the government sector and to take other complementary opportunities from outside the sector” (AGEST, 2007).

This theme also included references to ‘quality service’, or the ability, quality and dedication of the staff and management. For example:

“Lastly, but importantly, I also take this opportunity to express my appreciation to management and staff throughout all CSRF offices for their continued dedication in maintaining high standards of service to the Fund’s members.” (Catholic Superannuation and Retirement Fund, 2007).

The Profile theme covers statements which relate to the fund as an entity, including mention of experience (years of operation), size, strength, and growth. These statements often included statistics such as total assets under management, number of members, and contributions received. For example:

“After rapid growth last year, we received another record level of contributions and rollovers into the fund this year. The fund has $2.5 billion in assets and over 165,000 members at 30 June 2007.” (AGEST, 2007).

This theme also included examples of fund branding using the fund logo or image, such as:

“QIEC Super is proud to be an Industry Fund. Being an Industry Super fund means we value our members and operate on a not-for-profit basis which means all our profit goes straight back to the members.” (QIEC Super, 2008).

Specific references to the products and services offered by the Fund, especially new services, were coded to the Products and Services theme. This included references to insurance; the range of investment options offered; financial planning; member seminars; and website improvements. For example:

“Fund members now have an array of new choices such as personal online access, an insurance
upgrade, additional contribution methods and new investment options” (ACCSF, 2007). The Commitment theme captures statements which could broadly be characterized as conveying a ‘trust us’ message from the management or trustees. These statements were strong expressions of the Trustees’ dedication or commitment to members, usually expressed in personal terms. For example:
“Our motto ‘working just for you’ provides a clear focus.” (AustralianSuper, 2007); and
“Your superannuation investment is important to our team here at BT as well. We continue to work hard throughout the year to drive strong returns on your investment” (BT Lifetime Super, 2007); or
“You can be assured that your Trustee is dedicated to producing very competitive returns for you whilst minimizing the risk profile” (Managed Australian Retirement Fund, 2007).
Specific mention of trust was also included here, for example:
“It’s good to know that workers in the electrical contracting and communications industries can now access high quality financial advice from a trusted source like CONNECT.” (Connect Superannuation Plan, 2008).
References to where members’ had identified their trust through a survey were coded to ‘Member Satisfaction’ rather than Commitment.

III.2.3 Member
The third category of themes covers those messages aimed more specifically at the member. Statements which could be characterized as general Advice and Education were common. At one end of this theme was advice such as:
“The introduction of Simpler Super has made it very important to provide your tax file number (TFN) to the Fund. If the Fund does not have your TFN, your benefit may be taxed at a higher rate of tax” (First State Superannuation Scheme, 2007). Other statements could be characterized as more avuncular. For example:
“Compare super with the value of your house: quarterly price statistics you’ll read in newspapers show that house values move up and down. I wonder if you’ll see too many people rushing to sell their home just because prices in their suburb have dropped marginally!” (ACCSF, 2007), and
“Stick to your strategy and make sure you get advice from your financial adviser to confirm you are invested in the right areas of the market, given your needs and the type of investor you are” (BT Lifetime Super, 2008).
Some statements were less qualified and more direct including:
“When markets turnaround from a low point, they often do so with amazing speed. We can never know what the year ahead will bring, but we do know that if you are out of the market you risk missing out on the best returns.” (Asset Super, 2008), and
“The reality is that the market volatility we’ve seen over the last 12 months or so, looks set to continue into next year, and this will inevitably have an effect on investment returns in the short term, we have every reason to believe the outlook for the market is positive and we are confident that it will bounce back over the next 12 to 24 months.” (BT Lifetime Super, 2008)
Related to the Advice & Education theme were statements which could be distinguished by an empathetic tone, expressing an understanding of members’ feelings and circumstances, or their concerns about and confidence in their superannuation, again usually in personal terms. Such statements of Empathy included general messages of reassurance such as: “don’t panic”, “no need to worry”, or “it’s only natural to be concerned”. For example:
“However, we understand that some members, especially those close to retirement, may be
anxious about this year’s returns.” (Construction & Building Unions Superannuation (Cbus, 2008).

Other statements, while empathetic, were more dismissive such as:
“The important thing to remember is that super is a long-term investment, so worrying about short term volatility makes no sense” (ACCSF, 2007).

Some statements were congratulatory and complementary such as:
“So, well done to those of you who have been chipping in extra” (WALGSP, 2007); and
“Cbus appreciates that you have worked hard to build your super.” (Cbus, 2008)

The Satisfaction theme within the Member category of themes is the complement of the ‘Commitment’ theme within the Fund category. These statements highlighted positive member feedback from member survey, and made general reference to members’ satisfaction, confidence or trust. Examples include:
“Our member survey in late 2007 found that nine out of 10 members were satisfied with the OSF and trust that their super is ‘in good hands’.” (Officers’ Superannuation Fund, 2008), and
“During their research, Forethought found AUSCOAL Super members have the highest member satisfaction score in superannuation compared to all other super funds they have researched…” (AUSCOAL Superannuation Fund, 2008).

The Goals and Objectives theme is used where there was specific reference to members’ investment objectives, financial or retirement goals, retirement lifestyle, retirement security, ensuring adequate income in retirement, or maximizing benefits and savings for retirement. Examples include
“We have now established a strong foundation to meet these challenges and move forward with our objective of maximizing the retirement savings of our members.” (AustralianSuper, 2007), and
“The service is staffed by highly qualified salaried people who are committed to helping members maximize the amount of money they have to spend in retirement.” (Cbus, 2007).

This theme excludes general or vague references to long-term objectives and goals, which were coded to ‘Advice & Education’.

III.2.4 General
The final theme category is used for very general statements which in some cases were procedural. Such an example is:
“In this year’s report, you will find information on the Fund’s performance, financial statements, investment options, Trustee details, a message from the management and news on new products and services” (LUCRF, 2007).

Also included were general introductory statements, for example:
“I am pleased to present your Annual Report for the financial year ended 30 June 2007” (ING Masterfund, 2007); and concluding statements

IV. Results and Analysis
Chairperson or Chairperson/CEO statements are clearly viewed as important by superannuation funds as, in all but a handful of cases, they were placed as the first message that members read in the annual report. Of the sample of 81 funds and 162 reports, most often the statement appears on the second page of the report, with only one outlier case where the statement was closer to the end of the report.

The length and number of words in the statements is summarized in Table 3. There is a considerable spread in terms of word count and length (number of paragraphs) within each fund
classification. For example, in Retail funds the largest statement by word count has eight times as many words as the smallest in 2007. Corporate funds have the most words and paragraphs in both years, with Retail funds the least. The statements in Corporate funds, and to a lesser degree Retail funds, were longer in 2008, whereas Public Sector funds were shorter.

### Table 3 Word and Paragraph Count

<table>
<thead>
<tr>
<th>Fund Classification</th>
<th>Words</th>
<th>Paragraphs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2007</td>
<td>2008</td>
</tr>
<tr>
<td>Corporate Funds (n=8)</td>
<td>Mean</td>
<td>Max</td>
</tr>
<tr>
<td></td>
<td>664</td>
<td>1567</td>
</tr>
<tr>
<td>Industry Funds (n=42)</td>
<td>519</td>
<td>1036</td>
</tr>
<tr>
<td>Public Sector Funds (n=13)</td>
<td>633</td>
<td>1970</td>
</tr>
<tr>
<td>Retail Funds (n=18)</td>
<td>439</td>
<td>1216</td>
</tr>
<tr>
<td>Total (n=81)</td>
<td>534</td>
<td>1970</td>
</tr>
</tbody>
</table>

IV.1 Word Frequency

A word frequency analysis was conducted within each theme, rather than the whole statement. The corporate report literature (e.g., Smith & Taffler, 1992a) leads us to expect that:

- Bad news will be associated with an obfuscation of message and declining readability;
- Bigger companies will be associated with higher readability from their wish to signal their superior financial performance.

There are several confounding factors, however, that may distinguish the corporate narrative’s literature and the current analysis. Firstly, in the studies that have focused on corporate statements, analysis of readability is linked to financial performance measures including return on equity and likelihood of default. In the context of superannuation funds it is less clear what this measure of performance should be. APRA (2008, p.3) has stated it “considers that fund net ROAs are the best starting point from which to assess trustee performance” though this is a disputed measure (Cheong and Zurbruegg, 2008).

A second confounding factor for the current analysis is that all funds experienced lower returns in 2008. Unlike in the corporate analyses, where there is typically a distribution of good and bad performance, all funds had poor returns in 2008. There is, however, still a distribution of poor returns, and it is true that while a fund had poor returns they will have performed better (or worse) than other funds.

This suggests a performance measure which focuses on the returns earned on individual investment options may be useful. However, not all members are in the same investment option. It is true, though, that the majority of members remain in fund default investment options which suggests that as an indicator of overall fund performance this default return is a possible measure. The difference or turnaround in default returns, or the percentage drop in returns in 2008, is one measure of the magnitude of the turnaround in performance and provides some measure of the context for a fund’s statement. This measure, summarized for the sample in Table 4, ignores the risk of the default options but it provides some order of magnitude of the change in fortune of the funds.
Table 4 Change in Returns for Fund Default Options

<table>
<thead>
<tr>
<th>Fund type</th>
<th>Change in ROR (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Max</td>
<td>Min</td>
<td></td>
</tr>
<tr>
<td>Corporate Funds (n=8)</td>
<td>-147.65</td>
<td>-116.98</td>
<td>-180.58</td>
<td></td>
</tr>
<tr>
<td>Industry Funds (n=42)</td>
<td>-137.35</td>
<td>-110.99</td>
<td>-110.99</td>
<td></td>
</tr>
<tr>
<td>Public Sector funds (n=13)</td>
<td>-129.76</td>
<td>-84.95</td>
<td>-156.75</td>
<td></td>
</tr>
<tr>
<td>Retail Funds (n=14)</td>
<td>-173.56</td>
<td>-155.77</td>
<td>-200.67</td>
<td></td>
</tr>
<tr>
<td>Total (n=77)</td>
<td>-143.72</td>
<td>-84.95</td>
<td>-200.67</td>
<td></td>
</tr>
</tbody>
</table>

1 A return for both years could not be identified for four funds in the sample

Before analyzing readability and performance, there is some word frequency evidence of obfuscation in the reporting of bad news. Some funds appear to use language which reflects their reluctance to employ words which convey weakness in bad news periods. Thus where ‘strong’ is used to describe performance in good years, ‘negative’ rather than ‘weak’ is used to describe poor performance. There is also evidence of a renewed emphasis on superannuation as a ‘long-term’ investment in bad times (e.g., 2008), even though this emphasis was not apparent in 2007.

IV.2 Readability

The readability analysis in this paper focuses on the FLESCH scores, summarized in Table 5. The readability scores are surprisingly consistent, both between years and between fund type. The accounting literature suggests scores will be lower when bad news is disclosed; this is not the case with the superannuation statements. Indeed mean readability scores are marginally higher in 2008. The improvement is most marked for Retail funds.

There is no significant correlation between the change in the rate of return measure, presented in Table 4, and the 2008 FLESCH scores. There is, therefore, no evidence to support the signaling hypothesis. The results suggest a weak inverse relationship between fund size (asset base) and readability. The combination of high assets and low FLESCH scores suggests that the statements of smaller funds have higher readability, though the evidence is more suggestive of no relationship.

Table 5 FLESCH Readability Scores

<table>
<thead>
<tr>
<th>Fund type</th>
<th>Mean FLESCH Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2007</td>
</tr>
<tr>
<td>Corporate Funds (n=8)</td>
<td>37</td>
</tr>
<tr>
<td>Industry Funds (n=42)</td>
<td>40</td>
</tr>
<tr>
<td>Public Sector funds (n=13)</td>
<td>39</td>
</tr>
<tr>
<td>Retail Funds (n=14)</td>
<td>34</td>
</tr>
</tbody>
</table>

We have no reason to suspect that there will be readability differences between funds, though, there are some marginal differences in practice and the differences are minor when compared with the overall spread of scores recorded (20 to 60) which is beyond the range reported by Smith and Taffler (1992a) for corresponding accounting narratives. However, there does not appear to be any strong evidence of underlying characteristics (e.g., fund size or type) driving the differences, other than the writing style of the author.

IV.3 Themes

The average number of themes covered in the statements is eight, which varies little across the two years. To assess the importance given to a theme, a measure of its coverage in each statement was made. Each paragraph was coded to at least one theme. Given that the number of
paragraphs varied across funds, relying on the number of coded paragraphs alone would provide an inadequate measure of coverage. Therefore both word counts and number of paragraphs were utilized. In the simplest case if one theme is coded to a single paragraph, the coverage percentage is calculated as the word count for the respective paragraph divided by the total word count for the statement. In the case where multiple themes were coded to a paragraph, coverage was calculated by dividing the paragraph word count among the number of themes for the paragraph. For example, if three themes were coded to a paragraph with a total of 60 words, each theme would be assessed as contributing 20 words to the total word count of the statement.

The average coverage devoted to each theme and across all funds in both years is indicated in Table 6. In both years the two themes which had the largest coverage were “Operations, Goals and Objectives” and “Fund Services”. There was a degree of consistency in the level of coverage for three other themes across the years: “Investment Performance”; “Investment Strategy”; and “General”. The aggregate level coverage of “Investment Performance” appears marginally higher in 2008.

<table>
<thead>
<tr>
<th>Table 6 Theme Coverage</th>
<th>Mean theme coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theme category</td>
<td>2007</td>
</tr>
<tr>
<td><strong>External References</strong></td>
<td></td>
</tr>
<tr>
<td>Super Environment</td>
<td>12.66</td>
</tr>
<tr>
<td>Market Performance</td>
<td>5.30</td>
</tr>
<tr>
<td><strong>Fund</strong></td>
<td></td>
</tr>
<tr>
<td>Operations Goals &amp; Objectives</td>
<td>15.99</td>
</tr>
<tr>
<td>Investment Strategy</td>
<td>4.67</td>
</tr>
<tr>
<td>Investment Performance</td>
<td>9.61</td>
</tr>
<tr>
<td>Fund Services</td>
<td>12.94</td>
</tr>
<tr>
<td>Fund Profile</td>
<td>7.86</td>
</tr>
<tr>
<td>Fees</td>
<td>1.84</td>
</tr>
<tr>
<td>External Endorsement</td>
<td>2.49</td>
</tr>
<tr>
<td>Commitment</td>
<td>12.50</td>
</tr>
<tr>
<td>Board</td>
<td>5.37</td>
</tr>
<tr>
<td><strong>Member</strong></td>
<td></td>
</tr>
<tr>
<td>Satisfaction</td>
<td>0.51</td>
</tr>
<tr>
<td>Goals &amp; Objectives</td>
<td>2.70</td>
</tr>
<tr>
<td>Empathy</td>
<td>0.61</td>
</tr>
<tr>
<td>Advice &amp; Education</td>
<td>6.99</td>
</tr>
<tr>
<td><strong>General</strong></td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>9.21</td>
</tr>
</tbody>
</table>
While there is consistency in a number of themes as discussed, there are areas of noticeable difference in coverage between the years. In 2007 the significance and impact of the federal government’s “Better Super” changes is reflected in the large coverage of the “Super Environment” theme relative to 2008. This is common across each fund type. Similarly the larger “Fund Profile” coverage in 2007 is true for each fund type, with the exception of Corporate funds which had a lower coverage in 2007.

Looking at the themes in Table 6, three that had much larger coverage in 2008 stand out: “Market Performance”, “Advice and Education”, and “Empathy”. This increased coverage by all fund types is understandable given the significant events of 2007/08. The increase in coverage of these themes collectively is in keeping with the superannuation trust structure and is a demonstration of the fund expressing an awareness of issues of concern for members. A feature of these themes in 2008 was the emphasis placed on superannuation being a long-term investment; this was less evident in 2007.

IV.4 Performance Attribution

In terms of “Market Performance”, statements in 2007 describing performance typically use “strong” (for example: “delivered strong returns”) or “strength” (for example: “strength of the share market”). In 2008 the description of performance had changed most commonly to it being “negative” or showing “volatility”, not to “weak” or “poor”. These are not equivalents with an apparent preference in 2008 for what can be classified as more dispassionate words.

This raises the issue of whether there is a degree of selective attribution occurring. Supporting this view is the fact that 15 funds in 2008 made reference to how external markets “affected” returns. Similarly the use of “impact”, as in:

“The impact of the share market correction is now being felt by all large Australian super funds” (Mastersuper 2008).

These words were used in 2007 but to a lesser extent and in all but one case these were references to early signs of share market falls after the 2006/07 financial year end. In all but one case in 2007, and in all cases in 2008, “impact” was used as a means of explaining lower returns, not as a means of describing good returns.

Further supporting an attribution case is that there is no increased coverage given to the discussion of the investment performance of the fund itself, as there is for market performance. The question is, if the market performance is worthy of discussion in terms of its relevance for considering the 2008 fund investment performance, isn’t it also equally as worthy of discussion in 2007?

This is not to say that all reports display attempts at selective attribution. A relatively smaller number though acknowledged the link to good returns. For example:

“And the good news for investors is that the strength in the local market over the last 12 months has fed through into the performance of many of our superannuation investment options” (BT Lifetime Super, 2007), and

Self-attribution biases have long been documented (see Miller and Ross (1975) for an early critical discussion). More recently, Choi and Dong (2007) provide evidence of self-attribution bias in the behavior of mutual fund managers who they argue “tend to take too much credit for successes but too little blame for failures” (Choi and Dong, 2007, p.1). This is not to suggest an overly manipulative view of statement framing. It is more to suggest consistency of message would more likely make it easier for the 2008 message of the GFC to be delivered and accepted. To some extent the die is cast for the funds. They are vehicles for taking the contributions and investing on members’ behalf. The broad asset allocation decision will determine the vast
majority of their investment portfolio’s returns variability. Some will have different views as to the role of active and passive management, or the importance of the various services they offer members. Some may indeed have skills that consistently add value for their members. But 2008 highlights that while funds choose their own asset allocation, they don’t influence the various markets’ performance, rather they are influenced by them. Identifying this more consistently in good and bad return periods may be a better message approach.

IV.5 Benchmark Time Period
It has been suggested that in 2008 statements there was an increased emphasis on superannuation as a long-term investment. Corporations Regulations (reg 7.9.37(1)(j) require funds to publish average five-year return data in the annual report, though it doesn’t prescribe where. In 2008 ASIC pursued means of having funds provide more personalized five-year average return data in the member periodic statement as a means of “improving investor engagement with the long-term nature of their investment in super” (ASIC, 2008a). ASIC further suggested “it is more useful to convey such information with periodic statements than in annual reports” (ASIC, 2008b).

A check of the time periods used supports a definite change in language. Using the “Investment Performance” and “Market Performance” themes a word count was undertaken of the period of reference in both years, excluding reference to the current year. In 2007 there were 16 references to a five year investment period in the two themes. In 2008 there were 49. In 2007 there were five references to 10 or 20 year periods, whereas in 2008 there were 13.

IV.6 Advice & Education
In view of the increased coverage of the “Advice and Education” theme in 2008 it is interesting to note that ASIC recently introduced a class order relief (CO 09/210) allowing superannuation fund trustees to provide limited personal advice without relying on compliance with the suitability rule of section 945A of the Corporations Act. The relief aims “to improve the access of working Australians to advice about their super fund” (ASIC, 2009, RG 200.7). Within the “Advice and Education” theme, some statements fit neatly within ASIC’s definition of factual information as being “objectively ascertainable – its truth or accuracy cannot reasonably be questioned” (ASIC, 2007, RG 36.21).

Clearly no statements in the report offer personal advice as the statements are part of a report to all members. Many statements in the reports, 2008 in particular, provided what could be classified as opinion, with an arguable case that they carry implicit recommendations, which may be perceived as beyond general advice. The following two examples, made in different years, provide illustrations of this. The first example is more akin to general advice, whereas the second is closer to opinion and implicit recommendation.

“As the value of your account gets larger, it’s increasingly important to make sure you’re invested in an option that’s right for you. So it’s worth taking the time to review your investment choice.” (Asset Super, 2007).

“We can never know what the year ahead will bring, but we do know that if you are out of the market you risk missing out on the best returns.” (Asset Super, 2008).

IV.7 Style
The variation in the overall narrative style of fund statements is a further point of difference with corporate annual reports. Corporate reports have become more formulaic and ultimately less informative over the past two decades. Some fund statements share this and are very technical. These more technical statements are predominantly from Retail fund reports which appear quite different in intent and intended audience. This technical focus is reflected in lower FLESCH
scores previously discussed. The majority of funds however can be characterized as less formal with many presenting a more “folksy” style. The following examples illustrate the formal/informal contrast. The first is a more formal statement:

"Figures as at the end of June 2008 show that in less than 9 months, the new TasPension product has attracted 46 new members with assets totaling $5.7 million. The Tasplan Board are extremely pleased with this higher than anticipated take up and look forward to further growth in 2009." (Tasplan, 2008).

An example of an informal style providing the same information is:

"The Fund has again performed well this financial year with earnings of 13.3% and we now have over $66 million in assets under management and over 11,000 members nationally. You think children grow fast, what about this Fund!" (ACCSF, 2007).

In some instances this folksy style provides some interesting language such as the following example:

“The US sub–prime mortgage woes and credit crunch crisis have been a thorn in everyone’s side. Although, luckily, the Fund didn’t have much direct exposure. And, because these issues are likely to be around for the foreseeable future, markets will continue to be volatile.” (WALGSP, 2008) (underlining added).

This statement demonstrates some dangers inherent in the folksy approach. It is interesting to consider what role members consider luck plays in the management of their fund. Should members also consider it luck if fund returns are higher?

IV.8 Theme Grouping and Readability

The previous analysis has examined how much coverage was given to themes within the statements. It was noted that more than one theme could be coded to a single paragraph. Readability is enhanced when a paragraph is focused on a single topic or theme. Most often one or two themes were discussed in a paragraph, with an overall average of 1.5. In some cases, however, a single paragraph covered up to five different themes, detracting from the overall readability.

In 2007 the “Investment performance” theme was discussed more frequently by itself than “Market Performance”. “Investment performance” was also less frequently discussed with other themes, likewise “Market performance”. In 2008 there was quite a marked change. “Market Performance” was now more frequently discussed by itself compared with “Investment performance”. Further, the “Investment performance” theme was more frequently discussed with another theme in 2008. The fund’s investment performance was discussed in context of the market performance, while discussing an external endorsement, or with “Advice and Education” which generally provides a historical context. This context may all be relevant in 2008, but it is equally relevant in 2007. If this is a message that needs delivering, it may be better accepted if it appears in good and bad periods.

V. Discussion and Recommendations

Superannuation represents an increasingly important financial asset for an increasing number of Australian workers. Since the introduction of the Superannuation Guarantee in 1992 most superannuation fund members have only experienced negative annual returns in 2002. The negative returns in 2008 and into 2009 were larger and lasted longer. Any market downturn brings concern. In 2008, however, in addition to the negative returns, the member experience was also framed by institutional collapse within capital markets, even if this was U.S. and
European based. It was a self-evidently bad experience and presented a challenge to funds in what they communicated with their members. This paper has investigated how funds utilized one component of their communications to members. It is acknowledged that funds utilize other forms of communications including member magazines, newsletters, letters, and multimedia on their website. However, the chairperson’s statement is clearly an important communication. This is reflected in how it is positioned in the fund’s annual report. Further, the analysis of these statements allows consideration by funds as to whether the statement is consistent with the overall annual report and overarching communications strategy messages.

V.1 Fund Statement Readability
A methodology for analyzing these statements is available from the corporate annual report literature. Research in this area suggests unclear statements can be predictive of poor performance, and narratives may be constructed to obscure poor performance. A failure to establish a trusted relationship through the use of misleading or unclear statements may render statements ineffectual means of delivering messages. Whilst sharing much in structure with the annual reports of corporations, a fundamental difference is superannuation funds report to members, not shareholders. Further, in a majority of cases, they are compulsory members.

Analysis of a sample of 81 statements has identified many positive features. An analysis of readability suggests they improved when moving from a period describing the good news of 2007 to the bad news of 2008, contrary to evidence from corporate statement analysis. It remains to be established whether increased readability is associated with any increase in understanding of the statements.

An analysis of the major themes and coverage of these themes, in the statements identifies dominant messages. Three areas stand out with increased coverage in 2008: “Advice and Education”, “Empathy”, and “Market Performance”. The first two are indicative of the different nature of fund statements when compared with corporate narratives and is more reflective of the “vulnerable member” rather than “competent” investor model (Donald, 2008).

The increased coverage of the market experience through 2008 is expected but its context provides a richer story. The performance of external markets had as big an influence on fund returns in 2007 as they did in 2008. In 2008 statements though, the discussion of the markets was substantially larger, earlier, and more in combination with discussion of the fund’s own performance than it was in 2007. An increased coverage to discussion of funds’ own “Investment Performance” was restricted to Industry funds in 2008. Further, in 2008 the terminology used in describing performance was more dispassionate and emphasized how the market “affected” or “impacted” the fund’s return which was largely absent in 2007. Finally, the emphasis given to framing superannuation as a long-term investment supported by the increased description and use of longer benchmark periods is much higher in 2008.

V.2 Greater Consistency
Two overriding messages were delivered in the 2008 statements: superannuation should be considered as a long-term investment; and financial markets have a big impact on the fund’s investment options. It is hard to argue with these statements. However, following from this is that these views equally apply in good as in bad times. By being selective about emphasizing them the funds potentially reduce the potency of the message. It is therefore recommended that consistency of themes or message needs to be given greater consideration when framing statements. In particular, where funds seek to use statements to address the “vulnerable member”. It has been noted that there was an increase in the coverage of “Advice and
Education” themes in 2008. These themes can be undermined by opportunistic annual report statements which lack consistency. Consistent messages will strengthen advice and education messages and provide a basis for more avuncular statements suggesting that “worrying about short term volatility makes no sense”. Short term volatility in annual statement message may in fact be a contributor to this worry.

Notes
1. The implication here is not that the global financial crisis is now a historical event and thus over. Rather, the project will examine 2008 as a sample of what is collectively referred to as the global financial crisis.
2. Those born before 1st July 1960 can access at age 55 which increases one year, each year, until those born from 1st July 1964 who can access at a preservation age of 60 years.

References
Australian Bureau of Statistics (2009), Employment arrangements, retirement and superannuation, Australia, Apr to Jul 2007 (re-issue). Cat. No. 6361.0. Table 22, 26.
Beal, D. (2008), Superannuation and retirement income planning, John Wiley and Sons, Milton, Qld
OECD (2009), OECD Private Pensions Outlook 2008. OECD.


Appendix : List of Funds in Sample

<table>
<thead>
<tr>
<th>Corporate Funds</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAG &amp; NRMA Superannuation Plan</td>
</tr>
<tr>
<td>MasterSuper</td>
</tr>
<tr>
<td>Officers’ Superannuation Fund</td>
</tr>
<tr>
<td>Qantas Superannuation Plan</td>
</tr>
<tr>
<td>The Sisters of Mercy Staff Superannuation Scheme</td>
</tr>
<tr>
<td>Telstra Super Pty Ltd</td>
</tr>
<tr>
<td>Unilever Super</td>
</tr>
<tr>
<td>Water Corporation Superannuation Plan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Industry/Public Offer Funds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Government Employees Superannuation Trust [AGEST]</td>
</tr>
<tr>
<td>Australian Superannuation Savings Employment Trust [Asset Super]</td>
</tr>
<tr>
<td>Auscoal Superannuation Fund</td>
</tr>
<tr>
<td>The Allied Unions Superannuation Trust (Queensland) [AUST (Q)]</td>
</tr>
<tr>
<td>AustralianSuper</td>
</tr>
<tr>
<td>Building Unions Superannuation Scheme (Queensland) [BUSS (Q)]</td>
</tr>
<tr>
<td>Care Super</td>
</tr>
<tr>
<td>Construction &amp; Building Unions Superannuation [Cbus]</td>
</tr>
<tr>
<td>Club Plus Superannuation Scheme</td>
</tr>
<tr>
<td>Connect Superannuation Plan</td>
</tr>
<tr>
<td>Catholic Superannuation Fund</td>
</tr>
<tr>
<td>Catholic Superannuation and Retirement Fund</td>
</tr>
<tr>
<td>Cuesuper</td>
</tr>
<tr>
<td>Equipuper</td>
</tr>
<tr>
<td>Electricity Supply Industry Superannuation Fund (Qld) [ESI Super]</td>
</tr>
<tr>
<td>First Super</td>
</tr>
<tr>
<td>Health Super Fund</td>
</tr>
<tr>
<td>Health Employees Superannuation Trust Australia [HESTA]</td>
</tr>
<tr>
<td>Health Industry Plan</td>
</tr>
<tr>
<td>HOSTPLUS Superannuation Fund</td>
</tr>
<tr>
<td>Intrust Super Fund</td>
</tr>
<tr>
<td>Media Super</td>
</tr>
<tr>
<td>Iegalsuper</td>
</tr>
<tr>
<td>Labour Union Co-Operative Retirement Fund [LUCRF]</td>
</tr>
<tr>
<td>Meat Industry Employees’ Superannuation Fund [MIESF]</td>
</tr>
<tr>
<td>MTAA Superannuation Fund</td>
</tr>
<tr>
<td>Nationwide Superannuation Fund</td>
</tr>
<tr>
<td>National Catholic Superannuation Fund [NCSF]</td>
</tr>
<tr>
<td>New South Wales Electrical Superannuation Scheme [NESS]</td>
</tr>
<tr>
<td>Non Government Schools Superannuation Fund [NGS Super]</td>
</tr>
<tr>
<td>OAMPS Super Fund</td>
</tr>
<tr>
<td>Queensland Independent Education &amp; Care Superannuation Trust [QIEC Super]</td>
</tr>
<tr>
<td>Rei Super</td>
</tr>
<tr>
<td>Retail Employees Superannuation Trust [REST]</td>
</tr>
<tr>
<td>Statewide Superannuation Trust</td>
</tr>
<tr>
<td>Stevedoring Employees Retirement Fund [SERF]</td>
</tr>
<tr>
<td>Sunsuper Superannuation Fund</td>
</tr>
<tr>
<td>Tasplan Superannuation Fund</td>
</tr>
<tr>
<td>TWU Superannuation Fund</td>
</tr>
<tr>
<td>UniSuper</td>
</tr>
<tr>
<td>Vision Superannuation Fund</td>
</tr>
</tbody>
</table>

**Public Sector Funds**
- AvSuper Pty Ltd
- Brisbane City Council Superannuation Plan [City Super]
- CSS Fund
- Emergency Services Superannuation Scheme [ESSSuper]
- First State Superannuation Scheme
- GESB Superannuation Fund
- LG Super
- Local Government Superannuation Scheme
- Military Superannuation & Benefits Fund No 1
- State Public Sector Superannuation Scheme [Qsuper]
- Retirement Benefits Fund
- Southern State Superannuation Scheme [Triple S]
- WA Local Government Superannuation Plan [WALGSP]

**Retail/Public Offer/Master Trust Funds**
- [Accountants Super Fund] Division of Professional Associations Superannuation Fund
- Australian Child Care Super Fund [ACCSF]
- Astarra Superannuation Plan
- The Bendigo Superannuation Plan
- BT Lifetime Super
- FSP Super Fund
- ING Masterfund
- Map Superannuation Plan
- Managed Australian Retirement Fund
- Mercer Super Trust
- The Universal Super Scheme
- Perpetual Investor Choice Retirement Fund
- Russell Supersolution Master Trust
- Skandia Global Retirement Solutions
- Spectrum Super
- Summit Master Trust Personal Superannuation & Pension Fund
- Macquarie Superannuation Plan
- Netwealth Superannuation Master Fund
Two Cultures Shaping the Corporate Governance Debate

Donald G. Margotta*
Associate Professor of Finance
Finance and Insurance Group
College of Business Administration
Northeastern University
Boston. MA 032115
d.margotta@neu.edu

An early working paper version of this paper was presented at the Northeastern University College of Business Seminar Series in 2008.
Abstract
This paper discusses an alternative to the conventional finance framework for analyzing corporate governance problems in the modern corporation. That framework, going back to at least 1932, suggests that governance problems arise because managers control the corporation and may use their positions to enrich themselves at the expense of shareholders who are not able to effectively monitor them.

Alternatively, this paper suggests that some governance problems arise because major shareholders think differently about certain issues than do corporate managers, and these differences in thinking arise from their different “cultures.” The paper focuses on institutional shareholders and discusses how their culture, rooted in finance theory, may conflict with corporate managers whose cultures are based largely on legal theory.

Several examples illustrate how shareholders and managers may come to legitimate, but opposing views on certain corporate governance issues, and the paper stresses that these opposing views do not assume that either side acts with bad motives.

I. Introduction
The English physicist and novelist C.P. Snow stimulated wide discussion in 1961 when he suggested that people trained in the liberal arts thought about social and scientific issues in fundamentally different ways than those trained in the natural sciences. He said of this (Snow, 1961): “It was through living among these groups ... that I got occupied with the problem of what ....I christened to myself as the ‘two cultures.’ For constantly I felt I was moving among two groups - comparable in intelligence ... who had almost ceased to communicate at all ...”

This paper suggests that a similar observation can be made in the corporate governance area where one culture is represented by institutional shareholders, many of whose managers are trained in portfolio management and modern finance theory. The other is represented by corporate managers and directors who work in an environment shaped by corporate management and legal theory. Although these two have not “ceased to communicate” their communication on certain corporate governance issues is often adversarial. This paper examines why that might be.

While much of the finance literature assumes that the sometimes adversarial communications arises from what Jensen and Meckling (1976) refer to as “agency” issues, this paper suggests that another reason is that finance theory and legal theory sometimes conflict, leading individuals on either side to think differently about those issues. Understanding these differences and the conflicts they produce leads to suggestions of how the differences may be accommodated and the conflicts minimized.

This is a timely issue as activist institutional investors in the U.S. have become increasingly vocal in calling for governance changes in their portfolio companies, and as institutional activism spreads to other countries as well.

Differentiating National Cultures from Institutional Cultures
To see how a two cultures approach might apply in the corporate governance area it is important first to differentiate the conflicts under discussion here from those described in other literature.

8 Lublin, Joann and Sara Calian, “Activist Pension Funds Create Alliance Across Atlantic to Press Lackluster Firms,” Wall Street Journal, November 23, 1998. This article focuses on an alliance between the California Public Employees Retirement System, the largest public pension fund in the U.S., and Hermes Pensions Management, Britain’s largest pension manager. In addition, in “Samurai v. Shareholders, The Economist, February 16, 2008, page 75, describes a dispute between Sapporo an American investment firm Steel Partners and states that it is “only one of many rumbling disputes between Japanese companies and foreign activist investors.”
First, the proposition that different national cultures and legal systems give rise to different views on corporate governance has been understood for some time. For example, there is an extensive literature on corporate governance practices in countries such as England (Miller, 1998), Germany (Gordon, 1998), and Russia (Black, 2001), among others. However, while these studies address cultures in different countries, they do not address the differences in cultures of institutional investors and of corporate managements.

Second, to be sure, conflicts between investors and managements have been studied extensively. In fact, much of the corporate governance literature deals with such conflicts. That literature primarily suggests that the source of the conflicts lies in management dereliction of duty towards shareholders. Jensen and Meckling (1976) refer to these as “agency” issues or conflicts since they assume that managers are “agents” of shareholders and when managers are seen to be acting in their own interests rather than in the interests of shareholders they call that an agency conflict. That analysis essentially assumes bad motives on the part of managers and while such motives may certainly exist, this paper focuses on a different source of conflict, that arising from the differences in cultures of shareholders and management, cultures shaped by sometimes conflicting theories, not on bad motives, malfeasance or dereliction of duty on either side.

Defining Corporate Governance

The term “corporate governance” has come to have different meanings and it is therefore important to stress the difference between what has become the common usage of the term and the meaning used in this paper. Common usage, as noted earlier, deals with management malfeasance, i.e. with monitoring the actions of corporate officials to uncover or prevent illegal behavior. The Sarbanes-Oxley Act of 2002, for example, contains the most comprehensive set of corporate governance regulations passed in the United States and was a direct result of illegal corporate activity resulting in the bankruptcies of Enron and WorldCom in 2001 and 2002. Such activities pose important governance issues and have been the subject of much of the recent corporate governance literature.

In contrast, however, this paper deals with corporate governance issues involving decisions where no illegal activity is involved, such as when a company adopts a strategy that is legal, but contrary to the wishes of its largest shareholders. A company’s adoption of a poison pill takeover defense is an example that will be discussed in greater detail later in this paper. The governance question in such cases concerns who has, and who should have, the power to decide an issue when there is simple disagreement between shareholders and management over the efficacy of a particular strategy. A related question concerns why large shareholders in particular sometimes disagree with corporate managements on such issues.

Differentiating between monitoring for illegal activity and monitoring legal but controversial actions is important because this paper deals specifically with conflicts between large institutional investors and corporate managers and acknowledges that institutional investors may well be excellent monitors of corporate managements regarding malfeasance. However, the paper suggests that institutional investors may be inappropriate monitors on questions dealing with certain strategic issues. The difference is that in dealing with management malfeasance the interests of institutional investors in preventing illegal behavior is congruent with the interests of the corporation and its other shareholders, but in strategic issues their interests may conflict with those of the corporation and its other shareholders, especially with those shareholders who are undiversified, a point that will be developed further.

There are, of course, investors other than institutions who may have their own distinct cultures and conflicts with the corporation, such as corporate acquirers, risk arbitrageurs, high frequency
traders, and short sellers, but the focus here is on institutional shareholders for three reasons: their diversification is key to the discussion, they are the largest owners of corporate stock, and they are leaders in urging changes in corporate governance practices.

**Diversification of Institutional Portfolios is Key**

Diversification of institutional investors is key because it changes the way they view risk, and that differentiates them from corporate managers and undiversified shareholders. Modern Portfolio Theory (MPT), an underlying principle of modern finance theory, shows that diversified portfolios, such as those held by most institutional investors, are the most efficient ones, and through diversification investors can eliminate virtually all company specific risk. This insight, a breakthrough in finance theory, was developed by Harry Markowitz (1952) in a paper for which he was awarded the 1990 Nobel Prize in Economic Sciences, a prize he shared with co-recipients Merton Miller and William Sharpe. The impact of MPT is underscored by the fact that it has lead to indexing of many stock funds, meaning that these funds simply buy shares in companies that reflect a diversified market index. For example, Carleton et al. (1998) find that TIAA-CREF, the largest private pension fund in the U.S., indexes 80% of its domestic equity portfolio.

While MPT has created great benefits for shareholders it has also had unintended consequences in the area of corporate governance by creating potential conflicts over certain governance issues between diversified shareholders and the corporation because, by eliminating company specific business risk as an area of concern, diversified shareholders become more concerned with the performance of a portfolio of stocks than with individual stocks in the portfolio. Corporate managers, in contrast, are concerned with company specific risk because they have a legal duty to act in the interests of one corporation, not in the interests of a portfolio of stocks. As noted by Block et al. (1987), at least 30 states have codified their fiduciary duty statutes for directors: “Typical of the language in these statutes is para. 8.30 (a) of the Model Business Corporation Act: A director shall discharge his duties as a director... in a manner he reasonably believes to be in the best interests of the corporation.”

On most issues, of course, what is good for an individual corporation is also good for a portfolio containing that individual company, in which case there are no conflicts of interest. Monitoring a company to prevent illegal behavior is an example. However, this paper discusses decisions which directors believe are in the best interests of a corporation but might not be seen as being in the best interests of a diversified portfolio containing that company’s stock. These conflicts will be illustrated later with examples involving takeover defenses, bankruptcy, and product strategies.

**Changes in the Berle and Means Model**

Before discussing specific conflicts it will be helpful to examine the evolution of the relationship between stockholders and the corporation that led to the current situation, especially the changes observed in 1932 by Berle and Means and then the changes brought about through increased ownership of stock by institutional investors, especially since the 1980’s, and the diversification of those institutional owners.

This examination starts with Berle and Means’ “The Modern Corporation and Private Property,” one of the most cited works in the corporate governance literature (Berle and Means, 1932). The authors documented how ownership of major corporations was shifting away from individual owner-managers, i.e. owners who also managed their companies, to ownership by shareholders who, except for their ownership of shares, were no longer directly involved with the corporation as employees or managers. Simultaneously, control of the corporation was shifting to managers.
who were not major owners. They coined the often cited phrase, “the separation of ownership and control” to describe this phenomenon and that phrase has motivated much of the corporate governance research for more than 70 years.

Berle and Means wrote long before the development of modern finance and portfolio theory but it is helpful to put their description of changes in the ownership and control of corporations into a modern framework to understand the new forces at work. In that framework, modern finance theory looks at the total risk faced by corporations and investors as comprised of two parts, business risk and market risk.

Business risk involves risks such as competition, new products and technologies, and other factors specific to the corporation, while market risk involves factors external to the corporation such as interest rates and inflation rates. In early, pre-Berle, corporations and even young corporations today where founders still play a key management role and are large owners, the owners were also managers (owner-managers).

A representation of how such a Pre-Berle, pre-1932, corporation would look in a modern finance framework is shown in Figure 1 at the end of this paper. Figure 1 shows the undiversified owner-manager in a position of complete ownership and in complete control of the corporation. It was this owner-manager that Berle and Means said was being replaced in the modern corporation of 1932. Figure 1 also shows the corporation and the risks it faces from a modern finance viewpoint, i.e., business risk and market risk. In this model, the undiversified owner-manager bears both the business risk and the market risk faced by the corporation.

Figure 2 depicts the corporation as Berle and Means saw it in 1932 with shareholders separated from management but, again, in a modern finance theory framework, i.e. with the risks of the corporation separated into business risk and market risk. Although the authors noted that by 1932 shareholders were becoming increasingly separated from control of the corporation they did not discuss shareholder diversification since portfolio theory was not developed until the 1950’s. They focused instead on the separation of shareholders (ownership) from management (control) of the corporation, not on any other changes in the relationship of the shareholders to the company. Shareholders still had a relationship to the company, their financial interests were tied to those of the company, and they faced the same business risks and market risks as did the company.

The potential conflict of interest in this arrangement, as Berle observed, is that managers who control the company might use their control to usurp wealth from the owners and it is this conflict of interest that is the usual focus of corporate governance research, and similar to, as discussed earlier, what is now referred to as “agency” conflicts.

While these simple models do not describe all shareholders or all managements, they do illustrate the general evolution of the relationships between these two groups as viewed by Berle and Means around 1930.

Modern Portfolio Theory, as noted earlier, was developed in the 1950’s and its key insight was that investors can construct diversified portfolios to eliminate business risk. For example, a portfolio containing shares of Ford, General Motors, Nissan, DaimlerChrysler, and other auto companies can be immunized from competitive business risks such as a new model Ford taking away business from General Motors because the diversified portfolio contains shares of both Ford and General Motors. Similarly, if one company were acquired, as Chrysler was by Daimler-Benz in 1988, the diversified portfolio would not be eliminated from ownership in the automobile business since it would own shares in the remaining automobile companies, including the acquiring company. But the portfolio would still be exposed to market risks, such
as a depression that would hurt all automobile stocks.

Figure 3 illustrates what many large corporations face today, where their largest institutional shareholders, who typically own well over 50% of the company’s shares, are highly diversified (as they should be according to MPT) and have therefore eliminated business specific risk, leaving them exposed primarily to market risk, while the undiversified owners and managers continue to face both business risk and market risk.

To be sure, elimination of business specific risk is a powerful and beneficial result for diversified shareholders but, as Figure 3 illustrates, puts them in a different relationship to individual companies, to the managements of those companies, and to the companies’ undiversified shareholders.

Several observers have interpreted this development in different ways. Horrigan (1987), for example, commented on the position of the diversified investor from an ethical perspective. Applying Kant’s categorical imperative as his test of whether behavior is ethical, i.e. whether it would be acceptable if everyone acted in a particular way, he concludes that investors act unethically in adopting a portfolio perspective. His conclusion rests on an assumption that institutional investors have the necessary analytical skills to monitor individual corporations but cannot monitor effectively if their attention is on the portfolio. Deakin (2005), on the other hand, states that diversified institutional investors are more likely to engage individual companies to try to improve performance while Black (1992) suggests that institutional holders can play a constructive monitoring role. In reality, of course, there are institutions who are passive with respect to corporate governance initiatives, and others who do engage individual companies over specific governance issues.

From a “two cultures” perspective, there is no conflict of interest if diversified investors invest passively. Such investors simply buy stock and diversify their portfolios, as they should according to MPT. They are not seeking to control or influence the company. There are also no conflicts of interest when diversified shareholders seek changes that would benefit the corporation and its other shareholders. However, potential conflicts of interest arise between these investors and the corporation, and undiversified shareholders when they go further and seek to engage companies to effect certain other changes. The root of that potential conflict of interest lies in the different relationship of the institutional investor to the corporation created by diversification, as will be illustrated through the following examples.

**Examples: Conflicts between Diversified Investors and the Corporation**

**Takeover Defences**

Hostile takeovers and defenses against them provide an example of one of the most contentious topics on which institutional investors generally disagree with corporate managements. Institutional investors have welcomed takeovers and fought against efforts by companies to defend against them. In particular, the shareholder rights plan, more popularly known as a poison pill takeover defense, has been a lightning rod attracting protests and proxy proposals for their removal from institutional investors since the first case upholding them more than 20 years ago in Moran v. Household International.9 This is evidenced by numerous proxy proposals by institutional investors for companies to rescind the plans,10 while at the same time hundreds of

---

10 SharkRepellant.com is a research organization that tracks poison pill issues and it states the following regarding shareholder proposals to rescind poison pills: “With proxy season approaching, the debate between stockholder proponents of anti-poison pill proxy proposals and management will begin anew. Poison pill proposals have already been submitted for a vote at the 2005 annual meetings for Boeing Company, Morgan Stanley, McGraw-Hill
companies have adopted them as part of their takeover defense strategy. Similarly, institutional investors have generally opposed takeover defenses involving state takeover laws, while corporate managers have generally supported them, as have most state legislatures as evidenced by Boehm (1979) who states that at least 37 states in the U.S. having adopted some form of takeover law. Why do institutional investors and corporate managers view takeovers and takeover defenses so differently?

The conventional explanation from both groups is simple. Institutional investors and finance theorists generally view takeovers as positive events for target company shareholders since share prices of the target corporation rise when a bid is announced. They explain management opposition to such bids as an “agency” problem (Jensen and Meckling, 1976), where managers resist wealth gains for shareholders in order to foster their own interests, such as preserving their own jobs which are likely to be lost in the event of a takeover. A shareholder proxy proposal filed by The State of Wisconsin Investment Board (SWIB) opposing Champion International’s poison pill is typical of such proposals and states that the adoption of a pill lowers the probability that a company might be acquired and that a pill “effectively precludes a hostile takeover and thus allows management to take stockholders hostage.”

Managers, on the other hand, have viewed hostile offers negatively, citing reasons such as an offer “not being in the best interests of the corporation and its shareholders,” despite the premium price being offered, and frequently despite shareholder approval of the offer. They explain the enthusiasm of some shareholders for such offers as deriving from their having a “short-term” view. Champion responded to SWIB’s proxy proposal saying that a poison pill “enables the board to respond in an orderly and considered manner to an unsolicited bid. It puts the board in a better position to deflect unfair offers, such as coercive, partial, or two-tiered bids ... And it puts the board in a better position to negotiate a higher price for shareholders.” The latter point is supported by several studies which show that acquired companies having a poison pill defense get a higher takeover premium on average than those without a pill (Margotta, 1985). But a higher premium may not be persuasive from a portfolio perspective since a diversified owner’s portfolio could perform better if more companies were acquired, albeit at a lower premium. Further, since higher premiums are simply extracted from the acquiring company, and since it is likely that diversified owners also own the acquiring company shares, it may provide little net benefit if an acquiring company in their portfolio pays more for a target company in their portfolio since as owners of both they are simply paying themselves. In contrast, the higher premium would benefit the undiversified holders of the company’s stock.

The attempt by Microsoft to acquire Yahoo in 2008 provides an example of this. Yahoo directors rejected Microsoft’s initial offer of approximately $30 per share as inadequate. While there were efforts made to have Microsoft increase its offer, one analyst suggested that a majority of Yahoo’s largest institutional shareholders might not want a higher offer because they owned so many Microsoft shares that a higher offer by Microsoft would lower the value their Microsoft holdings. The analyst pointed out that 18 of Yahoo’s top 25 shareholders (42% of all Yahoo

---

Companies, PG&E Corporation, Maytag Corporation, Electronic Data Systems Corporation, Peoples Energy Corporation, ADC Telecommunications, and Allegheny Energy, Inc. The primary argument being made by the proponents of these proposals, and an argument generally supported by institutional investors based upon their proxy voting guidelines, is that poison pills entrench management and discourage takeover offers for the company.”

SharkRepellant.com website, Feb. 8, 2005

shares) actually owned more shares of Microsoft, which would tilt their interests toward Microsoft. “Of those top 18 shared stockholders, they hold 4.4 Microsoft shares for every one of Yahoo. If [Microsoft] were to raise the offer by $3 per share, and this negatively impacted its own shares by at least $0.68 … then these shareholders would be net losers,” He concluded that they may not want a bigger offer for that reason.”

Therefore, it is rational for diversified investors to oppose a takeover defense, especially if that defense might defeat an offer altogether. But it is equally rational from the perspective of the target company’s directors and management to embrace takeover defenses that enable them to negotiate a price that reflects their view of the value of the company, even if they defeat an offer which, although at a premium over the current market price, does not meet their estimate of the company’s value.

This last point can lead into an extensive discussion of efficient market theory (EMT), i.e. the idea that current prices are the best estimate of a company’s value. While that is an important discussion, it would be outside the scope of this paper. For the purposes here it is worth noting that finance theorists, and many large investors, generally accept EMT with the implication that managers should accept premium takeover offers and, more generally, should make decisions that maximize current share prices. This point is emphasized in Jensen and Smith (1984) where they state, “If capital markets are efficient, then the market value of the firm reflects the present value of the firm's expected future net cash flows, including cash flows from future investment opportunities. Thus the efficient markets hypothesis has several important implications for corporate finance. First, there is no ambiguity about the firm's objective function: managers should maximize the current market value of the firm.” But this only underscores the conflict with legal theory since there is no legal duty requiring managers to maximize current shares prices.

Clearly, every takeover would be different for each institutional holder and the effects described here would depend on the positions they held in each company, but it is clear that institutional investors and corporate managers may each have rational, yet different, perspectives on takeover defenses, without either having to impugn the other’s motives.

Bankruptcy

Bankruptcy provides another example where the two cultures might come to different, yet rational, views. Finance theory views bankruptcy as a company specific, diversifiable risk. Therefore, an institutional investor’s well diversified portfolio should not be greatly affected by the bankruptcy of one company in the portfolio. Fischel (1985) illustrates that point, saying: “The bankruptcy of an airline company, for example, might be a disaster for its employees and managers who lose their jobs but a matter of indifference to its investors who own shares in other airline companies that obtain the bankrupt company's routes.”

Fischel is correct from the perspective of the diversified stockholder. But bankruptcy is never a matter of indifference to a company's undiversified shareholders. Nor is it a matter of indifference to the company’s managers and directors, and not just because their jobs might be at stake. They have a fiduciary duty to act in the best interests of an individual corporation, not in the interests of a diversified portfolio of stocks, and they would be violating their duty under the

---


14 One exception to this under Delaware law is when a company is “for sale,” in which case the company enters what is referred to as “Revlon mode,” and in that case directors must accept the highest offer price. But as long as a company is not for sale directors are typically required to make decisions in what they believe is in the best interests of the corporation, even if the market and shareholders disagree.
law if they viewed bankruptcy as a matter of indifference. Therefore, directors might take actions to avoid bankruptcy in order to fulfill what they see as their duty to the corporation. One specific decision, for example, might be to carry less debt on its balance sheet to lessen the probability of bankruptcy. While that may seem prudent on the part of corporate directors, it might be viewed by diversified investors as being “underleveraged” and an effort by directors to avoid the stress of running the “tight ship” that a highly leveraged company would require. Jensen (1988), for example, argues that high levels of debt can create a beneficial control on managers by reducing the cash flow available for spending at their discretion.

As in the previous discussion of takeover defenses, institutional investors and corporate managers may each have different, and rational, perspectives on debt policy and bankruptcy arising from the respective positions and responsibilities, i.e. their cultures, without either having to impugn the other’s motives.

Competitive Strategy
Some strategic product decisions by a corporation are also likely to be viewed differently from the perspective of the two cultures. For example, in 2003 Kodak announced a decision to devote more resources to digital photography, and reduce its involvement in traditional photography. Explaining this decision, Kodak CEO Daniel Carp said, “We are acting in the knowledge that demand for traditional products is declining, especially in developed markets.”

Kodak shares dropped $4.84 per share, or 18% on the day following the announcement on trading volume of approximately 37 million shares, ten times the average daily volume, resulting in a one day loss of approximately $1.4 billion in market value based on 287 million shares outstanding. Several analysts and large investors questioned Kodak’s strategy when rivals such as Canon and Sony were said to be far ahead of Kodak in digital photography. While not suggesting that all of Kodak’s loss was due to portfolio rebalancing, this example illustrates how it would be rational from a diversified shareholders perspective to reduce their holdings of Kodak on this announcement. The entry of Kodak into this market adds little or no value to them since they are already invested in the digital market through their ownership of competitor shares. Having Kodak enter the market may even hurt their portfolio returns by adding another competitor. But for the Kodak Corporation and its undiversified shareholders such a strategic change is also rational. As with takeover defenses, and with debt policy, differences between institutional investors and management over strategic issues such as this can be understood as being rational as viewed from their respective cultures, but in conflict nevertheless.

The Conflicts Remain
Corporate governance conflicts have often been structured in terms of good vs. bad motives. Shareholders accuse management of acting in its own interests, while management accuses shareholders of being short term oriented. Each side sees its own views deriving from good motives, while deriding the others’ as deriving from bad motives. While bad motives on both sides are probably present in certain cases, it seems unrealistic to ascribe them too broadly to shareholders or managers, especially in light of the relatively transparent and highly litigious environment in which both exist. For example, on takeover issues, there may be offers where the good vs. bad model explains the behavior of some target corporation directors as well as the behavior of some finance

---

15 Wall Street Journal, September 26, 2003
professionals, but it is difficult to explain the actions of thousands of individuals on both sides engaged in hundreds of takeovers using such a simple model. It is especially difficult when one realizes that takeover decisions within target corporations are typically made by independent directors, and made by them with the advice of outside legal and financial counsel. In addition, every action taken by target directors is carefully scrutinized by the outside financial community, by the courts, and by the adversarial hostile acquirer. Any action taken by the target directors must be made in what they believe is the best interests of the corporation, and any violation of their obligation to do that exposes them to substantial legal liability.

It is equally difficult to explain the finance professional’s actions in terms of the good and bad motives model. Mangers of financial institutions such as pension funds and mutual funds have a fiduciary duty to act in what they believe to be the best interests of their beneficiaries, just as corporate directors and managers must act in what they believe to the best interests of the corporation. And fund managers also do not operate in a vacuum. Their decisions are subject to review, and are often a product of committee discussion, and are often made with outside counsel.

A two cultures approach attempts to set aside the good and bad motives model and instead explain how these two groups, acting in good faith, with due diligence, and with loyalty to their constituents, can come to legitimate but conflicting views on these governance issues. This approach also suggests ways to address the conflicts.

**Resolving the Conflict?**

One way is suggested by looking back at the efforts to resolve the initial corporate governance conflict described by Berle and Means, i.e. the separation of ownership and control. Those efforts were aimed at finding ways to motivate managers to think more like shareholders, for example by giving managers greater ownership stakes through stock grants and options. If managers had greater ownership stakes they would be expected to think and act more like owners.

Similarly, a two cultures approach suggests that significant corporate governance conflicts may arise between major stockholders and the corporation because diversification separates large shareholders from the business risk of the corporation which, it should be emphasized again, is desirable from an investment perspective.

If diversification is at the root of the conflicts of interest described here, several specific ways to resolve the conflicts arise, some of which are impractical but worth discussing if only to further illustrate the problem.

One way, for example, would be for governments to change the fiduciary rules for directors. Currently, directors are required by law to work in the interests “of the corporation.” But, as previously illustrated in Figure 3, the “interests of the corporation” are most closely tied to the interests of the undiversified owners. Should fiduciary duty statutes instead state that directors are to act in the “best interests of the diversified shareholders of the corporation?” That is probably impractical, but helps illustrate the conflicts within the modern corporation brought about by changes in finance theory.

A different approach to aligning the interests of institutional investors and management might be to have those investors who wish to exercise some influence on a corporation to restrict their holdings to those individual companies within specific industries, rather than diversify their holdings. That too would be impractical and bad investment advice since investors should be diversified.

A third approach could be to target specific issues where diversified stockholders might differ
with management and undiversified shareholders and limit voting on such issues to shareholders who do not own shares in competing companies. The U.S. Securities and Exchange Commission already makes judgments on certain corporate governance issues and allows companies to omit shareholder proposals from company proxy statements if it decides those proposals deal with operating decisions of the firm or director elections. So having the SEC make judgments on whether certain issues represent a conflict of interest for some shareholders would not be too great an extension of its current responsibilities.

While stockholders are currently limited in actions they can take to directly influence corporate policy, there are efforts underway to change that. Under U.S. federal regulations, for example, stockholders can introduce proxy proposals to eliminate poison pill takeover defenses but the vote on such proposals is non-binding on directors. However, because directors frequently do not act on these proposals, even when they receive a majority of votes cast, shareholders have adopted different tactics. One of these is reflected in Securities and Exchange Proposal 14a-11 introduced in 2003, which would give certain large shareholders the right to nominate directors under specific circumstances, one of which is triggered by director inaction on proxy proposals. This proposal has been supported by large institutional shareholders such as the California pension fund (CalPERS), and opposed by major business organizations. The director nomination proposal provides an example of a potential conflict of interest where the SEC could require shareholders who want to nominate directors certify that they do not own significant amounts of a competitor’s stock, thereby avoiding a situation where, for example, the largest shareholders of Ford end up nominating directors to General Motors, with all the conflicts of interest that might entail.

Recent state legislation in the U.S. also provides some guidance on this issue since many states have addressed the increasingly complex relationships between shareholders and the corporation, and between different groups of shareholders, although none have addressed diversified shareholders specifically. For example, more than 25 states have adopted control share takeover statutes which recognize the potential conflict between an acquiring person, who may be a company’s largest shareholder, and the rest of the shareholders. These statutes deny voting rights to anyone who goes over a certain ownership thresholds, typically starting with 20% ownership, unless the other shareholders vote to restore his voting rights. The constitutionality of such statutes was challenged and upheld by the U.S. Supreme Court.  

Other states, most notably Delaware, have adopted business combination statutes that prohibit a merger for a period of time if the acquiring company exceeds a threshold ownership level without the approval of the target company’s board of directors. In Delaware, the threshold ownership level is 15% of the shares outstanding and the moratorium on a merger is 3 years. This statute not only recognizes that an acquiring shareholder may have a conflict of interest with the other shareholders but it also recognizes others as having potential conflicts of interest. For example, the statute provides an exemption from the law if an acquiring person gets more than 85% of the outstanding shares in a tender offer, but it excludes from the calculation any shares held by the acquiring person and shares held by directors of the company, both of whom may have conflicts of interest with other shareholders on the question of a takeover.

Ohio provides another example in legislation it passed in 1990 that denies the right to vote in the control share vote to anyone who purchased more than $250,000 worth of stock after a takeover.

17 CTS Corp. v. Dynamics Corp. 107 S.Ct. 1637 (1987)
18 Delaware Code Section 203.
This statute is aimed at risk arbitrageurs and has special relevance for this discussion since the reasoning behind it is that arbitrageurs, although large shareholders, are short term holders and may have different interests than other shareholders. For example, in a typical stock transaction an arbitrageur will buy shares in a target corporation, while at the same time taking a short position in the acquiring company. He thereby locks in a profit determined in part by difference between the post-offer market price and the offer price since the market price immediately after an offer is announced is usually less than the offer price. By taking this position the arbitrageur, although a stockholder in the company, assumes no risk associated with the company. He is essentially an investor in the transaction, not in the company, and this legislation removes him from casting a vote on the future of the company.

Institutional investors are in a somewhat similar, although not entirely analogous situation. While institutions are longer term holders they, like arbitrageurs, assume little business risk associated with their holdings. Arbitrageurs eliminate business risk though their arbitrage position, while institutions eliminate it through their diversification. Arbitrageurs are investors in a transaction, and institutional shareholders are investors in the market, but neither assumes the business specific risk of an individual corporation.

Ohio implements its arbitrageur law by requiring persons voting in the control share election to certify that they do not fall into the class of shareholders prohibited from voting. States or federal regulators could similarly require that in certain votes where the interests of diversified holders might conflict with the interests of the corporation and its undiversified shareholders that they certify that they not significant shareholders of the acquiring company or of a competitor. This would essentially impose a duty of loyalty on shareholders in those situations, similar to the duty of loyalty that managers and directors have to the corporation.

Or, as suggested earlier, states could take an opposite tack and change their fiduciary duty statutes to relieve directors of their duty to the corporation and instead impose a duty of loyalty to the majority owners of the corporation, which will be the diversified institutional owners for most major corporations. While that is an unlikely outcome, it illustrates the central conflict of this discussion, which is that corporate managers and major corporate owners have different, and sometimes conflicting, responsibilities that lead to conflicts in how they view certain corporate governance issues.

Conclusion

This paper has attempted to first illustrate a change in the nature of ownership and control in the modern corporation brought about by the increased share ownership by highly diversified institutional investors. While Berle and Means documented a shift in ownership to smaller shareholders with little power to control managers, we are now witnessing a different shift of ownership to large institutions with the potential to influence management and corporate decisions. However, along with this, the paper also showed that because of the diversified portfolios controlled by these institutions their interests may not be the same as those of individual portfolio companies or of the undiversified owners of those companies. These conflicting interests help explain why diversified institutional investors might come to conclusions on key corporate governance issues which are opposed to conclusions reached by corporate managements. However, the common explanation in the finance literature for these opposing views has been that they arise from managements not acting in shareholder interests. Rather than attributing bad motives to either side, this paper suggests that each side, management

---

19 Ohio General Corporation Code, 1701.01 (CC)(1)(a)
and large investors, may come to rational, but different, conclusions given the theories which shape their respective cultures.

Scholars and practitioners tried to fix the agency related rift between managers and shareholders by recommending that managers be compensated more with stock, with the intention of motivating managers to think more like owners. The current environment may call for a different solution. If large shareholders seek to affect corporate decisions it may be desirable to have them certify that they are not significant owners of competitive corporations in order to vote on certain issues, such as those discussed in this paper. That would put such institutions in a difficult position because we know that from an investor perspective they are better off being diversified. But if they want to change their role from simply that of a passive investor to an activist investor, then it may be desirable that their interests be aligned with those of the corporation and its undiversified shareholders and meet the same test of loyalty to the corporation as is required of managers.

References

Deakin, Simon (2005), The Coming Transformation of Shareholder Value, Corporate Governance, V. 13, Number 1, January 2005
Gordon, Jeffrey, Deutsche Telekom, German Corporate Governance, And The Transition Costs Of Capitalism, Colum. Bus. L. Rev. 185.
Markowitz, Harry (1952), Portfolio Selection, Journal of Finance, 7, Number 1, 77-91.
Figure 1
Evolution of Corporate Control and Risk Taking
The corporation in the pre-Berle, Pre1930 era

![Diagram showing the evolution of corporate control and risk taking in the pre-Berle era.]

Figure 2
Evolution of Corporate Control and Risk Taking
The corporation in the post-Berle era, from the 1930’s through the mid-1980’s

![Diagram showing the evolution of corporate control and risk taking in the post-Berle era.]

Figure 3
Evolution of Corporate Control and Risk Taking: 1900’s to Present
The corporation in the post-MPT era, 1980’s to the Present

![Diagram showing the evolution of corporate control and risk taking in the post-MPT era.]
Revisiting Issues Involving Risk

David Eagle
Associate Professor of Finance
Eastern Washington University

Arsen Djatej*
Assistant Professor of Accounting
Eastern Washington University

Brian Grinder
Associate Professor of Finance
Eastern Washington University

Dean Kiefer
Associate Professor of Finance
Eastern Washington University

*corresponding author

Abstract
Most mortgages in the U.S. are either Adjustable Rate Mortgages (ARMs) or Fixed-Rate Mortgages (FRMs), both of which suffer from the tilt effect. Hence, changes in expected
inflation affects the affordability of ARMs and FRMs, potentially contributing to real estate bubbles such as occurred prior to the Financial Crisis of 2008. Also, FRMs expose both the borrower and lender to inflation risks while ARMs are an imperfect inflation hedge. On the other hand, Price-Level-Adjusted Mortgages (PLAMs) eliminate inflation risk, the tilt effect, and the effect of inflation on affordability. We find that the interest-rate risk and prepayment risk are no greater with PLAMs than with FRMs. We do, however, find a risk-sharing problem of PLAMs that could be dealt with.

Introduction
Since the Financial Crisis of 2008, the United States and many other countries have suffered the aftermath of one of the worst recessions since the 1930s Depression. This crisis was triggered by a complicated series of events involving real estate bubbles, mortgages, and derivative instruments supposedly set up to increase the liquidity of the mortgage markets. For more on the causes of this crisis, see Sanders (2008), Diamond & Rajan (2009), Shin (2009), Gramm (2009), Barth et al (2009), Blankenburg & Palma (2009), and LaCour-Little and Yang (2010).

In the United States, the vast majority of mortgages have been either Adjustable-Rate Mortgages (ARMs) or Fixed-Rate Mortgages (FRMs). Because the affordability of ARMs and FRMs are affected by the expected inflation rate, people can become confused and bid up housing prices irrationally when the expected inflation rate decreases as it did when our economies went through the Great Moderation. Thus, eliminating inflation as a factor in the affordability of mortgages should reduce the likelihood of future irrational housing bubbles.

Of particular concern is the tilt effect on both ARMs and FRMs, which causes the real value of the mortgage payment to be higher at the beginning of the mortgage compared to at maturity. The greater is the tilt effect, the higher is the initial real mortgage payment. The higher the initial real mortgage payment, the lower the real demand for housing and the lower the real price of housing. In summary, because of the tilt effect of ARMs and FRMs, the higher the expected inflation rate, the lower the real price of housing. Similarly, the lower the expected inflation rate, the higher the real price of housing.

However, Pareto efficiency requires that inflation not affect the real price of housing. Therefore, the effect of inflation on the real price of housing is an economic distortion. Brunnermeier and Julliard (2008) discuss how this distortion could potentially contribute to a real estate bubble like the one that preceded the Financial Crisis of 2008.

Because ARMs are based on nominal interest rates, the tilt effect also exists in ARMs. Price-Level-Adjusted Mortgages (PLAMs) eliminate the tilt effect by keeping the real mortgage payment constant over the life of the mortgage. For more about how PLAMs work, see Brueggeman and Fisher (2005).

When inflation rates fell as a result of the Great Moderation, interest in PLAMs by lenders, economists, and economic editors declined. Given the experience of the Financial Crisis of 2008, economists should again consider PLAMs because PLAMs are the only type of mortgage discussed in the previous literature that (i) offers true inflation protection and (ii) eliminates the tilt effect. This paper carefully articulates the distortionary roles of inflation on the housing market caused by ARMs and FRMs, discusses why ARMs are inefficient in its risk sharing.

---

20 The American Housing Survey estimated in 2006 that FRMs and ARMs accounted for 95.9% of the existing market, leaving 4.1% for other types of mortgages. However, based on Edgars and Bradley’s (2007) assessment of this Survey, these other types of mortgages probably accounted for much less of the mortgage originations in 2006, meaning the combined ratio for FRMs and ARMs is probably even higher.
properties, and discusses the risk-sharing properties of PLAMs. In particular, we find that the interest-rate risk and prepayment-risk of PLAMs should be less or at least no greater than FRMs, contrary to some of the preexisting literature. However, we do find one area of risk sharing where PLAMs are less than ideal, but that problem can be remedied.

Section I discusses the tilt effect and the distortional effect of inflation on the housing market. Section II discusses the interest-rate risk and prepayment risk of PLAMs. Section III discusses inflation and ideal risk sharing and applies this to the mortgage market. Section IV discusses the risk-sharing problems of ARMs as far as risk sharing. Section V summarizes the paper and discusses the one risk-sharing problem of PLAMs and how to fix it.

I. The Tilt Effect and Inflation’s Distortional Effect on Housing
For a $200,000 loan, Table I illustrates inflation’s effect on the tilt effect and the affordability of the initial mortgage payment for different expected inflation rates. We used the Fisher equation (See Chen, Liu, and Chang, 2005, and Rhodes, 2007) that \((1+i)=(1+r)(1+\pi)\) to determine the nominal interest rate from the 3% real interest rate and the expected inflation rate. Then based on this nominal interest rate, we determined the monthly mortgage payment.

\[
C/(1+\pi^t)\text{ where } C \text{ is the monthly nominal payment, } \pi \text{ is the expected inflation rate, and } t \text{ is the number of years from the mortgage’s inception.}
\]

The tilt effect is the decline in the real mortgage payment over time. A measure of the tilt effect is what we call the “tilt coefficient” which is the average slope of the real mortgage payment over 20 years as a percent of the original initial mortgage payment. This tilt coefficient is the average percent decline in the real mortgage payment per year.

It is important to distinguish inflation from the price level. The inflation rate is the percentage

<table>
<thead>
<tr>
<th>expected inflation</th>
<th>0%</th>
<th>2%</th>
<th>3%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominal rate</td>
<td>3.00%</td>
<td>5.06%</td>
<td>6.09%</td>
<td>8.15%</td>
<td>13.30%</td>
</tr>
<tr>
<td>rate per month</td>
<td>0.25%</td>
<td>0.41%</td>
<td>0.49%</td>
<td>0.66%</td>
<td>1.05%</td>
</tr>
<tr>
<td>monthly payment</td>
<td>$838.85</td>
<td>$1,067.09</td>
<td>$1,189.65</td>
<td>$1,448.13</td>
<td>$2,142.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>year</th>
<th>real payment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$838.85 $1,025.65 $1,121.36 $1,313.50 $1,770.75</td>
</tr>
<tr>
<td>4</td>
<td>$838.85 $985.83 $1,056.99 $1,191.38 $1,463.43</td>
</tr>
<tr>
<td>6</td>
<td>$838.85 $947.55 $996.31 $1,080.62 $1,209.44</td>
</tr>
<tr>
<td>8</td>
<td>$838.85 $910.75 $939.12 $980.15 $999.54</td>
</tr>
<tr>
<td>10</td>
<td>$838.85 $875.39 $885.21 $889.03 $826.07</td>
</tr>
<tr>
<td>12</td>
<td>$838.85 $841.39 $834.39 $806.37 $682.70</td>
</tr>
<tr>
<td>14</td>
<td>$838.85 $808.72 $786.50 $731.40 $564.21</td>
</tr>
<tr>
<td>16</td>
<td>$838.85 $777.32 $741.35 $663.41 $466.29</td>
</tr>
<tr>
<td>18</td>
<td>$838.85 $747.13 $698.79 $601.73 $385.37</td>
</tr>
<tr>
<td>20</td>
<td>$838.85 $718.12 $658.68 $545.79 $318.48</td>
</tr>
<tr>
<td>tilt coefficient</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Table 1. Real Mortgage Payment Over Time:
Example of a $200,000 loan with a real rate of 3%, compounded annually

Note: The nominal rate is also compounded annually computed as \((1+r)(1+\pi)\) where \(r\) is the real rate and \(\pi\) is the expected inflation rate. Therefore, to get the rate per month, we need to convert the nominal rate to an equivalent interest rate per month. Where \(i\) is the nominal annual rate, the equivalent rate per month equals \((1+i)^{(1/12)}-1\).

The Table computes the real monthly payment with the following formula: \(C/(1+\pi)^t\) where \(C\) is the monthly nominal payment, \(\pi\) is the expected inflation rate, and \(t\) is the number of years from the mortgage’s inception. The tilt effect is the decline in the real mortgage payment over time. A measure of the tilt effect is what we call the “tilt coefficient” which is the average slope of the real mortgage payment over 20 years as a percent of the original initial mortgage payment. This tilt coefficient is the average percent decline in the real mortgage payment per year.
change in the price level. In Table I the initial price level is the same no matter what the expected future inflation rate. Therefore, as Table I demonstrates, the greater the expected inflation, the greater the nominal interest rate, which means the greater is the tilt. The greater is the tilt, the greater is the initial real mortgage payment, and therefore the less the affordability is the first mortgage payment. The real initial monthly payment ranges from $838.85 for no inflation to over $2000 for 10% inflation, a more than doubling of the initial real monthly payment.

For most individuals, cash flow is important; thus, they usually judge what they can afford by that initial monthly payment. Also, mortgage underwriters’ have rules of thumb, such as “one’s total housing expense to one’s monthly gross income cannot exceed 28%,” that tie affordability to the initial monthly payment. Hence, a doubling of the initial monthly payment because expected inflation is 10% instead of 0% will sharply cut back on the demand for housing. Just as increases in expected inflation can reduce the demand for housing, decreasing in expected inflation can increase the demand for housing. By relying on ARMs and FRMs for our mortgages, we have structured the institution of mortgages such that changes in expected inflation will have substantial effects on the demand for housing.

In the early 1980s, when inflation expectations were high in the U.S. and Canada, a group of economists published forecasts on new mortgage designs. Many participants forecasted an end to fixed-payment, nominal interest rate mortgages. In particular, Hendershott and Villani (1983, p. 196-199) forecasted, “Mortgage contracts with nominal interest rates will be rare. … The early disappearance of fixed-rate mortgages is obvious.” Also, Car and Smith (1983, p. 226) likewise predicted that fixed-rate mortgages would be replaced primarily with PLAMs, although they also predicted ARMs would be used (which they were).

Nevertheless, Price-Level-Adjusted Mortgages (PLAMs) have rarely been used in the United States even though HUD has approved FHA insurance for PLAMs and the tax disadvantages of PLAMs have been eased (See Journal of Taxation, 1990). The lack of PLAMs in the marketplace has perplexed many financial economists (See Peek and Wilcox, 1991, and Weiner, 1983). In fact, even academic interest in PLAMs has waned since the 1990s.

The economists who made forecasts on mortgage design in 1980s did a rare, after-the-fact assessment of their forecasts twenty years later and acknowledged that their forecasts were way off the mark. Hendershott and Weicher (2002) noted that, “accurately forecasting inflation is crucial to forecasting housing markets.” However, what they had not forecasted was the lower and more stable inflation brought about by the Great Moderation in monetary policy in many countries including the U.S. and Canada. With inflation low and stable, the need for PLAMs to make mortgages more affordable was much less than in the past. As a result, lenders and economists turned their attention away from PLAMs and settled on the already established ARMs and FRMs to finance mortgages.

We can use Table I to illustrate how PLAMs affect the affordability of mortgages. PLAMs by design keep the real mortgage payment constant over the term of the mortgage. Hence, the real monthly payment on a PLAM for the example in Table I would be $838.85 regardless of expected inflation. When expected inflation is 10%, a PLAM could more than cut in half the initial mortgage payment, in this example from $2142.60 to $838.85. However, for the 2% expected inflation rate that is typical after the Great Moderation, the PLAM reduces the initial mortgage payments by about 20%, in this example from $1067.09 to $838.85. While the 20% is still significant, the “affordability need” for PLAMs is less acute with low inflation than with higher inflation.
That inflation has profound impacts on the housing markets is clear. However, Pareto efficiency requires that inflation have no impact on the real housing market when aggregate output of an economy remains the same. Hence, inflation’s impact on the real housing market is an economic distortion, an economic imperfection. For economic efficiency, we really need to inflation’s impact on the real demand for housing. To do so, we must eliminate the tilt effect on mortgages. By complacently sticking with ARMs and FRMs as the institutional setup for our financing of mortgages, we have kept our economies exposed to inflation’s distorting effect on the housing market. (Brunnermeier and Julliard, 2008) argue that inflation’s distorting effect could have contributed to the real estate bubble that preceded the Financial Crisis of 2008.

It is easy to see how the falling inflation scenario could lead to a real estate bubble. Inflation decreased, leading to lower nominal interest rates, making houses more affordable. This increased the demand for housing, causing house prices to increase. The continuing increase in house prices caused home owners to think those increases would continue, leading to more housing demand eventually leading to a real estate bubble.

To eliminate inflation’s distorting impact on the real housing market, we must eliminate mortgage tilt. In addition to PLAMs, the only other mortgage design in the current literature that can eliminate the tilt effect is the graduated-payment mortgage (GPM). However, the use of GPMs in the U.S. has been very limited.\(^\text{21}\) Also, for GPMs to eliminate the tilt effect, the rate of gradation must equal the expected inflation rate for the duration of the mortgage. However, GPMs as applied enable much steeper rates of gradation and as a result, the GPMs carry more credit risk and hence higher interest rates and are restricted to people who can document a high likelihood that their income will increase in the future. Also, because GPMs are based on a nominal interest rate, GPMs do not hedge against inflation risk and do experience negative amortization at the beginning of the mortgage.\(^\text{22}\)

II. Interest-rate Risk and Prepayment Risk of PLAMs vs. FRMs

In an attempt to explain why no market has developed for PLAMs, Leeds (1993) compares the levels of five different sources of risk between PLAMs and fixed-rate mortgages (FRMs). This paper focuses on what Leeds considers the two most significant risks -- prepayment risk and interest-rate risk (price risk). While several other financial economists have asserted without analysis that PLAMs would face less prepayment risk and less interest-rate risk than FRMs,\(^\text{23}\) Leeds’ more involved analysis (1989 and 1993) argues that PLAMs would have more prepayment risk and no less interest-rate risk than would FRMs. We disagree with Leeds' conclusions. This paper’s theoretical and empirical analysis indicates that PLAMs should not have a significantly greater amount of these risks than FRMs. Our analysis conflicts with Leeds’ analysis primarily because Leeds uses the interest-rate elasticity of perpetuity to approximate a mortgage’s elasticity. We, on the other hand, determine the exact interest-rate elasticity of a mortgage and find that PLAMs have significantly lower interest-rate elasticity than do FRMs. Using these interest-rate elasticities and some empirical results, we find that PLAMs do not have higher relative price risk, and may have lower real relative price risk than do FRMs. Because the

\(^{21}\) The American Housing Survey estimated in 2006 that GPMs amounted to 1.2% of existing mortgages. However, Edgars and Bradley (2007) note that this survey has had known problems, that most of these reported GPMs are either carryovers from previous years or misclassification because of the nature of the survey. From other sources, GPM originations have been quite limited.

\(^{22}\) PLAMs do experience nominal negative amortization, but not real negative amortization.

\(^{23}\) See Peek and Wilcox (1991, p. 52, p. 59, p. 60),
lower relative price risk that PLAMs should experience is substantially lower, PLAMs should also experience less prepayment risk as well. These findings that prepayment risk and interest-rate or price risk are no greater for PLAMs than for FRMs are important because they allow us to concentrate on other barriers to PLAMs that are more easily overcome. In particular, one of the other barriers that existed in the early 1990s was that other inflation-indexed financial instruments in the U.S. were rare. However, since 1997 the U.S. Treasury has been issuing TIPS, which are inflation-indexed bonds. With the existence of a major inflation-indexed security, funds specializing in inflation-indexed instruments have been already introduced, which then could lead to sources of funding for PLAMs.

A second issue working against PLAMs is that they adjust for inflation regardless the cause. While the literature on PLAMs has yet to discuss this issue, the wage indexation literature has. We will discuss this issue in the final section of the paper.

As Fischer (1984) states, ideally inflation indexing should only filter out the aggregate-demand-caused inflation, but leave aggregate-supply-caused inflation intact. In the final section of this paper, we will further address this issue.

The next subsection derives the interest-rate elasticity for both FRMs and PLAMs and then finds that PLAMs have significantly lower interest-rate elasticities than do FRMs. We then use the true interest-rate elasticities and some empirical analysis to show that PLAM’s should experience no more prepayment risk, less real relative price risk and less nominal relative price risk than do FRMs.

**The Interest-Rate Elasticities of FRMs and PLAMs:**

Assuming no prepayments, the (nominal) present value of the future payments of a FRM is the present value of its nominal mortgage payments (discounted by the current nominal interest rate):

\[
P = \sum_{r=1}^{n} \frac{C}{(1+i)^n} = C \frac{1-(1+i)^{-n}}{i} \tag{1}
\]

where \( P \) is the present value of the future payments of the FRM, \( C \) is the nominal mortgage payment, \( i \) is the nominal interest rate per month, and \( n \) is the number of months in the term of the mortgage. Differentiating (1) with respect to \( i \) and then dividing the result by equation (1) and multiplying by \( i \) gives the nominal interest rate elasticity of a FRM:

\[
\frac{dP}{di} = \frac{in}{(1+i)^n - 1} \tag{2}
\]

For a PLAM, the present value of the future payments and the interest-rate elasticity are the same as (1) and (2) except we discount by the real interest rate \( r \) instead of \( i \), \( C \) represents the real mortgage payment of the PLAM, and the result is the real interest-rate elasticity of a PLAM.

Define the interest elasticity function to be:

\[
e(r) \equiv -\frac{rn}{(1+r)^n - 1} \tag{3}
\]

where \( r \) here refers to the pertinent interest rate which would be a real rate for a PLAM and a nominal rate for a FRM. Differentiating with respect to \( r \) gives:

\[
e'(r) = -\frac{n((nr-1)(1+r)^n + 1)}{((1+r)^n - 1)^2 (1+r)^2} \tag{4}
\]
This derivative is negative for \( r > 0 \) since the numerator and denominator of the fraction in (4) are both positive when \( r > 0 \). Since \( e'(r) < 0 \) and since L'Hospital's rule (See Rubin, 1976, p. 109) shows that \( \lim e(r) < 0 \) as \( r \) approaches 0, it follows that \( e(r) < 0 \) for all \( r > 0 \). Furthermore, \( e(r) > -1 \) for all \( r > 0 \), since the fraction on the left in the brackets in (3) is positive when \( r > 0 \). In summary, \( e(r) \) is between -1 and 0 and it decreases as \( r \) increases. Therefore, since the real interest rate on a PLAM will be less than the interest rate on a FRM, the PLAM’s interest-rate elasticity will be closer to zero than the FRM’s interest-rate elasticity. This conclusion differs from Leeds’, because her perpetuity approximation led to an interest elasticity of -1, which didn’t vary with the interest rate.

While this mathematical argument shows that the interest-rate elasticity is less in absolute value for a PLAM than it is for a FRM, we need to investigate whether this difference is significant. Table II presents the interest-rate elasticities implied by (3) for different interest rates between 1% and 14%. For example, a 4%, 15-year PLAM would have an interest elasticity of -0.27, whereas an 8%, 15-year FRM would have an interest elasticity of -0.48. In this example, the FRM would have 1.8 times the interest elasticity of the PLAM, clearly a significant difference. For 30-year mortgages, the differential in interest elasticity is lower, but still significant. This differential increases as the nominal rate (and by implication the inflation rate) increases.

This subsection demonstrates that not only is the interest-rate sensitivity lower for PLAMs than for FRMs, it is substantially lower. The next subsection uses this result and empirical interest-rate volatility in the United Kingdom and the United States to assess the interest-rate risk and

<table>
<thead>
<tr>
<th>Annual interest rate</th>
<th>Monthly interest rate</th>
<th>Elasticity (15-year)</th>
<th>Elasticity (30-year)</th>
<th>FRM-PLAM ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td>14%</td>
<td>1.17%</td>
<td>-0.71</td>
<td>-0.94</td>
<td>2.6</td>
</tr>
<tr>
<td>13%</td>
<td>1.08%</td>
<td>-0.68</td>
<td>-0.92</td>
<td>2.5</td>
</tr>
<tr>
<td>12%</td>
<td>1.00%</td>
<td>-0.64</td>
<td>-0.90</td>
<td>2.4</td>
</tr>
<tr>
<td>11%</td>
<td>0.92%</td>
<td>-0.61</td>
<td>-0.87</td>
<td>2.2</td>
</tr>
<tr>
<td>10%</td>
<td>0.83%</td>
<td>-0.57</td>
<td>-0.84</td>
<td>2.1</td>
</tr>
<tr>
<td>9%</td>
<td>0.75%</td>
<td>-0.53</td>
<td>-0.80</td>
<td>1.9</td>
</tr>
<tr>
<td>8%</td>
<td>0.67%</td>
<td>-0.48</td>
<td>-0.76</td>
<td>1.8</td>
</tr>
<tr>
<td>7%</td>
<td>0.58%</td>
<td>-0.44</td>
<td>-0.71</td>
<td>1.6</td>
</tr>
<tr>
<td>6%</td>
<td>0.50%</td>
<td>-0.38</td>
<td>-0.64</td>
<td>1.4</td>
</tr>
<tr>
<td>5%</td>
<td>0.42%</td>
<td>-0.33</td>
<td>-0.57</td>
<td>1.2</td>
</tr>
<tr>
<td>4%</td>
<td>0.33%</td>
<td>-0.27</td>
<td>-0.48</td>
<td>1.1</td>
</tr>
<tr>
<td>3%</td>
<td>0.25%</td>
<td>-0.21</td>
<td>-0.38</td>
<td>1.0</td>
</tr>
<tr>
<td>2%</td>
<td>0.17%</td>
<td>-0.14</td>
<td>-0.27</td>
<td>0.9</td>
</tr>
<tr>
<td>1%</td>
<td>0.08%</td>
<td>-0.07</td>
<td>-0.14</td>
<td>0.8</td>
</tr>
</tbody>
</table>

NOTES: *This is the ratio of the interest-rate elasticities on a FRM at the particular interest rate relative to a 4% PLAM.

---

The denominator of the fraction in (3) is clearly positive when \( r > 0 \). For the numerator, define \( g(r) \equiv n ((nr - n)(1+r)^{n+1}) \). Note that \( g'(r) = n(n^2r(n+1)(1+r)^n+1) > 0 \) for all \( r > 0 \). Since \( g(0) = 0 \) and \( g'(r) > 0 \) for all \( r > 0 \) then \( g(r) > 0 \) for all \( r > 0 \).
Interest-Rate Risk and Prepayment Risk of PLAMs vs. FRMs:

Interest-rate risk is comprised of price risk and reinvestment risk. Price risk is the risk that the value of a security (e.g., mortgage) will change because of a change in interest rates. Price risk is relevant to investors who may need to sell the security before the security reaches maturity or to financial institutions whose sources of funds for investing in the security have maturities differing from the maturity of the security. We will focus our analysis, as did Leeds (1989 & 1993), on the price risk of PLAMs and FRMs due to changes in interest rates.

The results of the previous section have implications concerning the relative price risk and prepayment risk of PLAMs and FRMs. What is important to an investor is the relative price volatility. Since the interest rate elasticity is defined as \( \frac{dP}{P} \frac{i}{di} \), we can solve for \( dP/P \) to get:

\[
\frac{dP_{FRM}}{P_{FRM}} = e_{FRM} \frac{di}{i}.
\]

where \( e_{FRM} \) is the interest-rate elasticity of a FRM. If we take standard deviations (represented by \( \sigma \)) of both sides of (5) and treat the interest-rate elasticity as a constant,\(^{26}\) we get:

\[
\sigma \left( \frac{dP_{FRM}}{P_{FRM}} \right) = |e_{FRM}| \cdot \sigma \left( \frac{di}{i} \right)
\]

Similarly, for a PLAM,

\[
\frac{dP_{PLAM}}{P_{PLAM}} = e_{PLAM} \frac{dr}{r}
\]

\[
\sigma \left( \frac{dP_{PLAM}}{P_{PLAM}} \right) = |e_{PLAM}| \cdot \sigma \left( \frac{dr}{r} \right)
\]

where \( P_{PLAM} \) is the price of the PLAM.

To estimate the relative interest-rate volatilities of nominal and real interest rates, we used the constant-maturity series on nominal and real government bonds in the U.K. and the U.S. (In the U.K., the government bonds are called “gilts”.) The U.K. has issued inflation-linked gilts longer than the U.S. government has had its Treasury Inflation Protected Securities (TIPS). The Bank of England publishes the U.K.’s constant maturity series, which began in 1995. The U.S.’s constant-maturity series began in 2003, which we got from the FRED database from the Federal Reserve Bank of Saint Louis (2009).

Because PLAMs are not currently offered in the United States, we assumed the spread relationship between the 30-year mortgage rate and a Treasury rate would be the same for PLAMs as for FRMs. We used the following simple linear model of this spread:

\[
i_{FRM} = a + b \cdot i_{NG} + \varepsilon
\]

\(^{25}\)Reinvestment risk is the risk that the return on funds reinvested from the loan payments will differ from the original interest rate. Houston (1988) discusses the reinvestment risk of PLAMs and finds that PLAMs have less reinvestment risk than do FRMs.

\(^{26}\)In fact, the interest-rate elasticities are not constants; they vary with the interest rate as shown in equation (3). Therefore, technically speaking, we should keep the interest-rate elasticities inside of the standard-deviation operator. However, the inequality resulting from using the approximations in equations (6) and (8) would still result from a more rigorous, and complicated analysis.
where \( a \) and \( b \) are constants, \( \varepsilon \) is an error term, \( i_{\text{NG}} \) is the interest rate on a nominal government bond with a particular maturity. We assume that the interest rate on a PLAM will be:

\[
i_{\text{PLAM}} = a + b \cdot i_{\text{NG}} + \varepsilon
\]

(10)

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Correlation with MORTG</th>
<th>Constant</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9212</td>
<td>3.89%</td>
<td>0.8147</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(50.63)</td>
</tr>
<tr>
<td>3</td>
<td>0.9610</td>
<td>2.86%</td>
<td>0.9097</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(74.32)</td>
</tr>
<tr>
<td>5</td>
<td>0.9742</td>
<td>2.20%</td>
<td>0.9708</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(9.34)</td>
</tr>
<tr>
<td>7</td>
<td>0.9793</td>
<td>1.73%</td>
<td>1.0107</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(103.43)</td>
</tr>
<tr>
<td>10</td>
<td>0.9821</td>
<td>1.34%</td>
<td>1.0489</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(111.56)</td>
</tr>
<tr>
<td>20</td>
<td>0.9826</td>
<td>0.41%</td>
<td>1.1333</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(102.66)</td>
</tr>
</tbody>
</table>

Notes: The figure in parentheses under the coefficients are the t-statistics of the coefficients. Results reflect linear regressions of MORTG against GS1, GS3, GS5, GS7, GS10, and GS20 from April 1971 through June 2009. The degrees of freedom where 457 except for the regression involving GS20, where the degrees of freedom were 376 because of missing GS20 data between January 1987 through September 1993.

Table III: MORTG Regressed onto U.S. Treasury Constant-Maturity Series

where the coefficients \( a \) and \( b \) are the same as for a FRM of the same maturity. Table III presents the estimated coefficients of these models for the United States, where we regressed the 30-year monthly mortgage rate (Fred Database, 2009, MORTG) on the constant-maturity nominal U.S. Treasury interest rates. One important result is the correlation between MORTG and the U.S. Treasury series was greater the higher was the maturity of the U.S. Treasury series. We used the models of Table III to estimate the hypothetical nominal interest rates and hypothetical real interest rates on PLAMs using the nominal and real constant maturity U.S. Treasury series. (FRED database, 2009, GS5, GS7, GS10, GS20, FII5, FII7, FII10, FII20). With these hypothetical mortgage interest rates, we then produced the empirical results presented in Table IV, which we use to assess whether interest-rate risk (price risk) or prepayment risk will be greater for PLAMs than for FRMs. Table IV first shows the elasticities for both FRMs and PLAMs based on the maturity of the U.S. government securities we used. Secondly, the table presents the relative interest volatility (the standard deviations of \( \text{di}/i \) and \( \text{dr}/r \)). Thirdly, the table presents the relative price volatility (the standard deviations of \( \text{dP}/P \) for FRMs and PLAMs), which by equations (6) and (8) are determined by the product of the elasticity and the relative interest volatility. (Because the elasticities do change when interest rate changes, the relative price volatility will differ slightly from the product of the mean elasticity and the relative interest-rate volatility.

The comparison between the relative price volatilities between FRMs and PLAMs reveals that the relative price volatility of PLAMs would be significantly statistically greater than for FRMs.
if we relied on the 5-year or 7-year U.S. Treasury rates. However, as Table III points out the correlation between the 30-year and U.S. Treasury rates is greater for longer-term securities. Therefore, it would be more appropriate to rely on the long-term constant maturity series. For the relative price volatility estimated from the 10-year and 20-year U.S. Treasury rates, PLAMs would be less than for FRMs, although this difference is not statistically significant.

The data on constant maturities of inflation-indexed government bonds in the U.S. only exists since 2003 since the U.S. has only recently been issuing inflation-indexed bonds. The U.K., on the other hand, has been issuing inflation-indexed gilts since 1982. Table V presents a similar analysis to Table VI except for the U.K. Table VI’s results are much stronger in support of PLAMs than for FRMs as the relative price volatility of PLAMs is statistically significantly less than the relative price volatility of FRMs for all maturities.

Prepayment risk is very closely related to interest-rate risk. The more the interest rate falls, the greater is the likelihood that the mortgage holder will refinance and prepay the mortgage. Since PLAMs have less interest-rate risk than do FRMs or at least no greater risk than do, PLAMs should also experience less prepayment risk than will FRMs.

Leeds (1993) concluded that PLAMs would have more prepayment risk than would FRMs even though she had (erroneously) determined that the relative price risk of PLAMs would be the same as for FRMs. She reached this conclusion because she argued that the likelihood of refinancing would increase with the absolute price risk of the mortgage instrument. Leeds argued that the loan principal on PLAMs would be greater than that on FRMs because people could afford more expensive homes with PLAMs than with FRMs. However, under McCulloch’s (1986) underwriting recommendation of a maximum payment to income ratio of

<table>
<thead>
<tr>
<th>5-year maturity:</th>
<th>FRM</th>
<th>PLAM</th>
<th>% diff.</th>
<th>F-Stat</th>
<th>p-value</th>
<th>Test Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest elasticity</td>
<td>-0.21232</td>
<td>-0.12062</td>
<td>76%</td>
<td>5.86***</td>
<td>0.00%</td>
<td>Mean Diff</td>
</tr>
<tr>
<td>relative int. rate volatility</td>
<td>0.03101</td>
<td>0.04764</td>
<td>54%</td>
<td>2.36***</td>
<td>0.01%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>price volatility</td>
<td>0.00667</td>
<td>0.00541</td>
<td>23%</td>
<td>1.52**</td>
<td>3.54%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>int. rate volatility</td>
<td>0.00024</td>
<td>0.00019</td>
<td>30%</td>
<td>1.68**</td>
<td>1.25%</td>
<td>Var. Diff.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10-year maturity:</th>
<th>FRM</th>
<th>PLAM</th>
<th>% diff.</th>
<th>F-Stat</th>
<th>p-value</th>
<th>Test Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest elasticity</td>
<td>-0.36974</td>
<td>-0.21118</td>
<td>75%</td>
<td>7.64***</td>
<td>0.00%</td>
<td>Mean Diff</td>
</tr>
<tr>
<td>relative int. rate volatility</td>
<td>0.02816</td>
<td>0.03827</td>
<td>36%</td>
<td>1.85***</td>
<td>0.41%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>price volatility</td>
<td>0.01042</td>
<td>0.00756</td>
<td>38%</td>
<td>1.9***</td>
<td>0.29%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>int. rate volatility</td>
<td>0.00021</td>
<td>0.00014</td>
<td>52%</td>
<td>2.31***</td>
<td>0.02%</td>
<td>Var. Diff.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20-year maturity:</th>
<th>FRM</th>
<th>PLAM</th>
<th>% diff.</th>
<th>F-Stat</th>
<th>p-value</th>
<th>Test Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest elasticity</td>
<td>-0.51670</td>
<td>-0.31939</td>
<td>62%</td>
<td>7.44***</td>
<td>0.00%</td>
<td>Mean Diff</td>
</tr>
<tr>
<td>relative int. rate volatility</td>
<td>0.02391</td>
<td>0.03327</td>
<td>39%</td>
<td>1.94***</td>
<td>0.22%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>price volatility</td>
<td>0.01229</td>
<td>0.00974</td>
<td>26%</td>
<td>1.59**</td>
<td>2.25%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>int. rate volatility</td>
<td>0.00013</td>
<td>0.00009</td>
<td>43%</td>
<td>2.05***</td>
<td>0.10%</td>
<td>Var. Diff.</td>
</tr>
</tbody>
</table>

Table IV: Elasticity and Volatility Comparison Between PLAM and FRM based on constant maturity U.K. Treasury Series.
20% for a PLAM and 28% for a FRM, the maximum loan on a 30-year, 4% PLAM would be only approximately 10% more than the maximum loan on a 30-year, 8% FRM.\footnote{For example, the monthly maximum payment on a $3000 monthly income would be $600 for a PLAM and $840 for a FRM. Therefore, the maximum PLAM loan would be $125,677, which is 9.78% greater than the maximum FRM loan of $114,478.}

To the extent that refinancing costs are unaffected by the size of the loan principal, the absolute price risk is appropriate for analyzing prepayment risk. However, to the extent refinancing costs vary with the size of the loan principal, refinancing should vary with the relative price volatility of the mortgage. Since both absolute price risk and relative price risk are less with PLAMs than with FRMs, the prepayment risk on PLAMs should be lower than that on FRMs.

### The Pareto Ideal of Risk Sharing Between Borrowers and Lenders

Economists are well aware of the risk of unexpected inflation or deflation on borrowers and lenders when they use nominal loans like FRMs. While economists have called this risk “inflation risk,” the risk is in fact “price-level” risk not inflation risk. The distinction between “inflation risk” and “price-level risk” is extremely important in some contexts such as whether it is better for the central bank of a country to target inflation or the price level. The true economic risk is price-level risk, not inflation risk. As such, from this point in this paper, we will refer to this risk as “price-level risk,” despite the literature almost always referring to it as “inflation risk.”

This price-level risk is as follows: If the future price level is greater than the borrower and lender expected, then the real value of the nominal loan payment is less than expected, making the borrower better off at the expense of the lender. On the other hand, if the future price level is less than expected, then the real value of the nominal payment is higher than expected, making the lender better off at the expense of the borrower. If both the borrower and the lender are risk averse, then they both would be better off without this price-level risk. Therefore, if a mortgage is designed so that the real value of these payments are unaffected by changes in the price level, both the borrower and the lender are made better off. This is what PLAMs do by making the real value of the mortgage payment not only same regardless of the time of the payment, but also regardless of the price level.

Campbell and Cocco (2002, p. 22) analyze the welfare improving potential of PLAMs relative to FRMs. They find that PLAMs do lead to substantial welfare gains, half of which is due to eliminating the tilt effect, and half of which is due to eliminating price-level risk. It is important to note that GPMs potentially, although not in practice, can eliminate the tilt effect. However, GPMs do not eliminate the price-level risk. Thus, the potential welfare gains of GPMs is only half the potential welfare gains of PLAMs.

It is important to note that the logic that both borrowers and lenders gain when price-level risk is eliminated only applies when “everything else remains the same” including aggregate output. When aggregate output changes, things are more complicated as the reflections section of the paper discusses.

### III. Risk-Sharing Problems of ARMs

While Leeds (1993) compared PLAMs to FRMs, the mortgage markets in the U.S. and other countries have utilized adjusted rate mortgages (ARMs) as well as FRMs. In the U.S. the use of ARMs peaked in 1994 at 60 to 70 percent of all mortgage originations (Moench et al, 2010), although the average between 1995 and 2009 was 20 percent (Krainer, 2010). Canada, the UK, and Australia are more dependent on ARMs (Moench et al, 2010). In particular, in the UK, 65 percent of mortgages held between 2000 and 2002 were variable rate mortgages (Smith, 2005).
Many people think ARMs protect borrowers and lenders from inflation or price-level risk. To some extent, ARMs do reduce price-level risk. In particular, the Fisher equation tells us that the nominal interest rate approximately equals the real interest rate plus the expected inflation rate plus a premium for price-level risk.

If inflation increases, then in time the expected inflation will increase, which by the Fisher effect will increase the nominal interest rate. Under this circumstance, an ARM’s interest rate would increase as a result of higher expected inflation. On the other hand, if inflation decreases, then in time the expected inflation will decrease, which by the Fisher equation will decrease the nominal interest rate. Under this circumstance, an ARM’s interest rate would decrease as a result of

<table>
<thead>
<tr>
<th>5-year maturity:</th>
<th>FRM</th>
<th>PLAM</th>
<th>% diff.</th>
<th>F-Stat</th>
<th>p-value</th>
<th>Test Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest elasticity</td>
<td>-0.13719</td>
<td>-0.09124</td>
<td>50%</td>
<td>2.16**</td>
<td>3.91%</td>
<td>Mean Diff</td>
</tr>
<tr>
<td>relative int. rate volatility</td>
<td>0.05145</td>
<td>0.09046</td>
<td>76%</td>
<td>3.09***</td>
<td>0.00%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>price volatility</td>
<td>0.00628</td>
<td>0.00854</td>
<td>36%</td>
<td>1.85***</td>
<td>0.40%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>int. rate volatility</td>
<td>0.00022</td>
<td>0.00029</td>
<td>34%</td>
<td>1.81***</td>
<td>0.53%</td>
<td>Var. Diff.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7-year maturity:</th>
<th>FRM</th>
<th>PLAM</th>
<th>% diff.</th>
<th>F-Stat</th>
<th>p-value</th>
<th>Test Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest elasticity</td>
<td>-0.18618</td>
<td>-0.12017</td>
<td>55%</td>
<td>2.92***</td>
<td>0.57%</td>
<td>Mean Diff</td>
</tr>
<tr>
<td>relative int. rate volatility</td>
<td>0.05322</td>
<td>0.09067</td>
<td>70%</td>
<td>2.9***</td>
<td>0.00%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>price volatility</td>
<td>0.00869</td>
<td>0.01155</td>
<td>33%</td>
<td>1.76***</td>
<td>0.71%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>int. rate volatility</td>
<td>0.00022</td>
<td>0.00029</td>
<td>31%</td>
<td>1.71**</td>
<td>1.03%</td>
<td>Var. Diff.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10-year maturity:</th>
<th>FRM</th>
<th>PLAM</th>
<th>% diff.</th>
<th>F-Stat</th>
<th>p-value</th>
<th>Test Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest elasticity</td>
<td>-0.25995</td>
<td>-0.16250</td>
<td>60%</td>
<td>4.22***</td>
<td>0.01%</td>
<td>Mean Diff</td>
</tr>
<tr>
<td>relative int. rate volatility</td>
<td>0.05317</td>
<td>0.06510</td>
<td>22%</td>
<td>1.5**</td>
<td>3.98%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>price volatility</td>
<td>0.01204</td>
<td>0.01076</td>
<td>12%</td>
<td>1.25</td>
<td>16.44%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>int. rate volatility</td>
<td>0.00022</td>
<td>0.00019</td>
<td>15%</td>
<td>1.33</td>
<td>10.92%</td>
<td>Var. Diff.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20-year maturity:</th>
<th>FRM</th>
<th>PLAM</th>
<th>% diff.</th>
<th>F-Stat</th>
<th>p-value</th>
<th>Test Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest elasticity</td>
<td>-0.46433</td>
<td>-0.26431</td>
<td>76%</td>
<td>6.61***</td>
<td>0.00%</td>
<td>Mean Diff</td>
</tr>
<tr>
<td>relative int. rate volatility</td>
<td>0.52854</td>
<td>0.06358</td>
<td>731%</td>
<td>69.11***</td>
<td>0.00%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>price volatility</td>
<td>0.02091</td>
<td>0.01746</td>
<td>20%</td>
<td>1.43*</td>
<td>5.93%</td>
<td>Var. Diff.</td>
</tr>
<tr>
<td>int. rate volatility</td>
<td>0.00021</td>
<td>0.00016</td>
<td>28%</td>
<td>1.64**</td>
<td>1.60%</td>
<td>Var. Diff.</td>
</tr>
</tbody>
</table>

Notes: The table reports the mean interest rate elasticity and the standard deviations for the particular volatilities being measured. The F-statistic for the Mean Difference Test is $|\mu_1 - \mu_2|/\sqrt{S_1^2 + S_2^2}/n$ and the F-statistic for the variance difference test is the greater of $S_1^2/S_2^2$ or $S_2^2/S_1^2$. (See Freund, 1988). To help read the table, we bolded the font for the lower elasticity or volatility measure.

Table IV: Elasticity and Volatility Comparison Between PLAM and FRM based on constant maturity U.S. Treasury Series.
higher expected inflation. However, the inflation protection provided by ARMs is not a perfect hedge for the price-level risk of the ARM for several reasons. First, the interest rate in the Fisher equation responds to the expected inflation rate, not to the expected price-level per se. Depending on what has happened in the past with the price level, a greater or lower inflation rate could conceivably move the economy closer to the original expected price level. Second, it is expected inflation not the actual inflation that affects the nominal interest rate. A perfect hedge would adjust the mortgage’s payments to what happens with the actual price level. Third, the Fisher equation is a long-run relation; in the short run factors other than inflation can affect interest rates. For example, contractionary monetary policy will increase the nominal interest rate in the short run while decreasing the expected inflation rate in the long run. As a result, an ARM resetting during such a time would end up increasing its interest payments even though future inflation will actually decrease.

Prior to the 1980s, Barth et al (2008) report that the vast majority of mortgages were (i) written and held by savings and loan institutions and (ii) were traditional, fixed-rate, 30-year mortgages. However, the savings and loans funded these long-term mortgages with short-term deposits. Hence, when inflation increased in the 1970s, which led to higher nominal interest-rates, savings and loans started to lose money which led to the savings and loan crisis. This was an extreme example of higher inflation making the borrowers better off at the expense of the lenders. Since the savings and loan crisis in the 1980s, U.S. financial institutions are required to hedge their interest rate risk. One way of doing so is for them to issue ARMs instead of FRMs. This has been part of the reason for the increased use of ARMs in the U.S. However, ARMs transfer the interest-rate risk from the lender to the borrower. When increases in the nominal interest rate are due to factors other than inflation, the borrowers may suffer as a result. In particular, during the Financial Crisis of 2008, banks became more leery of lending to other banks causing LIBOR to increase. Many ARMs and most subprime mortgages were based on LIBOR. Hence, when these ARMs reset, the increased credit-risk premium in the interbank lending market caused the interest rates and the mortgage payments on these ARMs to increase (Schweitzer and Venkatu, 2009). Many borrowers were unable to meet these higher mortgage payments and defaulted on their obligations (See Knox, 2006).

IV. Summary and Reflections

While the U.S. mortgage markets were beginning to seriously consider PLAMs in the early 1980s when inflation was high, the lower and more stable inflation brought about by the Great Moderation, caused the markets to rely on the already established ARMs and FRMs as well. By doing so, the institutional structure of our mortgage markets is prone to changes in inflation, which leads to economic distortions. These distortions can lead to housing bubbles such as the one that preceded the Financial Crisis of 2008.

Economic efficiency requires that we eliminate the effect of inflation on our mortgage and housing markets. Inflation affects these markets and the borrower and lender in two ways: (i) the tilt effect and (ii) price-level risk. GPMs can potentially, but not in practice, eliminate the tilt effect, but it does not eliminate the price-level risk. ARMs can partially reduce price-level risk, but it does not eliminate that risk completely and it exposes borrowers to other risks that cause interest rates especially in the short run to change for reasons other than inflation. PLAMs, on the other hand, eliminate both the tilt effect and the price-level risk. By eliminating the tilt effect, PLAMs can reduce the initial mortgage payment by 20% relative to ARMs and FRMs.
While the U.S. has not seen much in the way of inflation-indexed mortgages, other counties with higher inflation have. For example, Brazil, Israel, Turkey (See Erol and Patel, 2007), and Mexico have utilized versions of inflation-indexed mortgages. Brazil has used PLAMs, whereas Mexico has used a Dual Index Mortgage which bases the interest rate for loan amortization on a variable interest rate or inflation, whereas it bases the monthly mortgage payment on a wage index. A problem with the Dual Index Mortgage is that at times the mortgages experience negative amortization. Brzeski et al (2009) discuss Poland’s use of a Dual Indexed Mortgage and how PLAMs may be a better alternative for the Polish mortgage market. On the other hand, in Mexico, Dale-Johnson and Towle (2002) report that a version of the Dual Index Mortgage that includes insurance against negative amortization “behaves like a PLAM but with reduced credit risk due to a lower probability of payment shock than is typically associated with PLAMs in Mexico.”

Given that PLAMs eliminate both the tilt effect and price-level risk, economists often are perplexed as to why more loans are not inflation indexed. One of the important implications of this paper’s findings is that we can no longer look at prepayment risk or interest-rate risk (price risk) as reasons why PLAMs have not been embraced in the markets. Hence, we must look elsewhere for the reasons for the absence of PLAMs in the U.S. marketplace. One reason for the lack of a PLAM market given by other financial economists (See, for example, Weiner, 1983) is that PLAMs and PLADs (Price-Level-Adjusted Deposits) would need to be created simultaneously in order that PLAMs have a source of inflation-indexed funds. With the recent U.S. Treasury’s program of issuing inflation-indexed bonds, some Price-Level-Adjusted Funds do now exist. The existence of these Price-Level-Adjusted Funds may then become a catalyst not only for funding for inflation-indexed bonds, but for funding inflation-indexed mortgages as well.

Perhaps, the reason financial institutions have not embraced PLAMs is the same reason firms’ interest in wage indexation diminished after the oil shocks in the 1970s (See Adolph & Wolfstetter (1991). As Grey (1976, 1978 and (Fischer, 1977) discuss, wage indexation does not efficiently handle real inflation shocks. It is also true, that other inflation-indexed contracts also do not efficiently handle real inflation shocks. Earlier in the paper, when we explained why borrowers lose (gain) while lenders gain (lose) when the price level is lower (higher) than expected, we stated that logic only applies when real aggregate output remains the same. When we have shocks to real output, inflation may be needed to appropriately distribute the impact of this real shock between borrowers and lenders. In other words, inflation (or deflation) may be needed in order for both borrowers and lenders to appropriately share in the real output shock. This is a point discussed by Eagle and Domian (2005) and has been recently been rediscovered by Koenig (2011).

Suppose real aggregate output falls by 1%. Other than storage, this would mean that aggregate consumption must fall by 1%. This means that average consumption must fall by 1%. If individual A’s consumption drops by less than 1%, then there must exist some individual B whose consumption drops by more than 1%. That may be appropriate if both A is more risk averse than B. However, if both A and B have the same level of relative risk aversion, then Pareto efficiency requires that A’s and B’s consumption must change drop by the same proportion. In particular, if both A and B have average relative risk aversion, then their consumption must drop proportionately to real aggregate output. The real value of a prearranged nominal payment behaves in this manner when the central bank of an economy targets nominal GDP. However, when the inflation is caused not by a supply shock, but by a demand shock, then
that inflation works against Pareto efficiency.

As (Fischer, 1984) states, ideally an indexation scheme should be one which adjusts for aggregate-demand-caused inflation, but leaves aggregate-supply-caused inflation. The quasi-real indexing designed by Eagle and Domain (1995) meets Fischer’s ideal with relative simplicity. Modifying PLAMs in this fashion is a worthwhile pursuit, but one which we will leave to future research (See, Eagle and Domian, 2011).

Shiller et al (2011) propose continuous workout mortgages (CWMs) to deal with situations where market conditions of the value of the house being less than the loan. In Mexico, they have had some insurance similar to CWMs. However, insurance has to cover such losses and that insurance adds to the expense of the mortgage. Barbatz (2004) reports, “The loans are relatively expensive – inflation plus 13.7% is no small amount.” Instead of continuous workout mortgages, Posner and Zingales (2009) recommend a legal reform that enable homeowners to reduce their principal in exchange for giving the mortgage holder some equity in the property. Shiller et al (2011) also are perplexed as to why so much of consumers’ equity in the United States is tied up in housing. However, Deng and Quigley (2007) find that there is convincing theoretical and empirical evidence that “households could be made much better off by a functioning market in house price derivatives.”

References


Deng, Y. and Quigley, J. (2007). “Index Revision, House Price Risk, And The Market For Home Price Derivative,” working paper no. w07-003 at the University of California,
Berkley


Representativeness Heuristic Can Cause Asset Price Underreaction to New Information in an Asset Market with Strategic Interactions

Guo Ying Luo
Department of Finance and Business Economics
DeGroote School of Business
McMaster University
1280 Main Street West, Hamilton, Ontario, Canada
luogy@mcmaster.ca

I gratefully acknowledge the comments and suggestions from the editors and anonymous reviewers. In addition, I would like to thank my colleagues from McMaster University for helpful comments and the Social Sciences and Humanities Research Council of Canada (SSHRC) for financial support.
Abstract
This paper presents a model of imperfect competition among rational traders, representativeness heuristic traders and noise traders in an asset market. Rational and heuristic traders both receive an informational signal about the asset's payoff before any trade takes place. In updating their beliefs about the asset's payoff, heuristic traders place too much weight on the current informational signal and not enough weight on their prior information. To maximize their expected profits, all rational and heuristic traders strategically submit their market orders to the market maker who sets the asset price equal to the expected asset's payoff conditional on the observed aggregate demand. This paper shows that representativeness heuristic can cause not just asset price overreaction to new information but can cause asset price underreaction to new information.

Keywords: representativeness heuristic, Asset Price Overreaction and Underreaction to New Information, Behavioral Model, Asset Market with Strategic Interaction

I. Introduction
The phenomena of asset price overreaction and underreaction to new information have been well documented in the literature. They are inconsistent with the efficient market theory, nor can they be explained by the existing asset pricing models. A few behavioral models resort to the psychological evidence (namely, behavioral biases) for explanations. For example, Daniel, Hirshleifer and Subrahmanyam (1998) uses self attribution bias to explain the causes of asset price underreaction and overreaction to new information. The conservatism bias and representativeness heuristic are used in the model of Barberis, Shleifer and Vishny (1998) to produce the asset price underreaction and overreaction to new information. Hong and Stein (1999) regards technical traders' extrapolation of the past pattern as the cause of asset overreaction to new information.

There are other explanations for the causes of the asset price underreaction and overreaction to new information. For example, Hirshleifer, Lim and Teoh (2011) use investors' limited attention to different types of information to explain the causes of both under- and over-reaction to different earnings components. Hong and Stein (1999) regards the asset price underreaction to new information as a result of the gradual diffusion of information across the population.

In spite of this, representativeness heuristic is popularly viewed as one cause of asset price overreaction to new information in the empirical literature. For example, Douks and McKnight (2005) and Jegadeesh and Titman (2001) empirically test and support the predictions of the models of Barberis, Shleifer and Vishny (1998), and Hong and Stein (1999).

Representativeness heuristic is one type of psychological behavioral bias documented by psychologists in their experiments (see Grether (1980), Kahneman and Tversky (1973), and Tversky and Kahneman (1974)). Tversky and Kahneman (1974) state "A person who follows this heuristic evaluates the probability of an uncertain event, or a sample, by the degree to which it (i) is similar in its essential properties to the parent population, (ii) reflects the salient features of the process by which it is generated" (p. 33). Hence, when forming posterior beliefs, people with representativeness heuristic place too much weight on their current information and too little weight on their prior knowledge. They underweight the base rate relative to Bayes rule. This is also referred as the base rate fallacy in the psychological literature (see Grether (1980), Kahneman and Tversky (1973), and Tversky and Kahneman (1974)). Also, technical traders' over-extrapolation of the past prices is another example of representativeness heuristic.

This paper examines the impact of representativeness heuristic on the asset price in a security
market with strategic interaction among traders. It proves that representativeness heuristic can cause asset price underreaction to new information. This result is in a sharp contrast with the literature attributing representativeness heuristic as the cause of asset price overreaction to new information. Note that this paper does not attempt to explain the empirical anomaly of asset price overreaction and underreaction to new information. Instead, it focuses on demonstrating that representativeness heuristic is capable of generating asset price underreaction to new information in an asset market with strategic interaction among traders.

Specifically, this paper, in the spirit of Kyle (1985), builds a one-period model of an asset market. The asset's payoff is unknown to all traders in the beginning of the period but traders receive an informational signal about the asset's payoff before any trade takes place. There are rational traders, heuristic traders and noise traders. Heuristic traders in this model place too much weight on the current informational signal and not enough weight on their prior information when updating their beliefs. Rational and heuristic traders are both risk neutral. Noise traders trade for their liquidity needs. Hence, their demand for the asset is assumed to be random. There is one market maker who supplies the liquidity to the market. The cost of doing so is assumed to be zero. To maximize their expected profits, all rational and heuristic traders strategically submit their market orders to the market maker. After observing the aggregate market orders of all traders, the market maker competitively sets the asset price equal to the expected asset's payoff conditional on the observed aggregate market orders. The market maker does not observe the informational signal about the asset's payoff.

In the equilibrium, the market order coming from each of rational and heuristic traders is generated from maximizing his expected profits given all others' equilibrium market orders and given the market maker's equilibrium pricing rule. Given the equilibrium market orders of all rational and heuristic traders, the asset price equals the expected asset's payoff conditional on the observed aggregate demand. The asset price in the equilibrium is shown to be a linear function of the aggregate demand. The coefficient of the aggregate demand is called the pricing parameter for the sake of following discussions. In addition, the equilibrium market order of each of rational and heuristic traders is also shown to be a linear function of the informational signal.

This paper finds that in the equilibrium, the asset price can overreact to some informational signals under certain model parameters conditions and can underreact to the same information or different ones under other model parameters conditions.

In this paper, the reason that the asset price underreacts or overreacts to some good news is as follows. Note that the change in the asset price resulting from the increase in the proportion of heuristic traders (i.e., the derivative of the asset price with respect to the proportion of heuristic traders) can be broken into two components. One component is the impact on the asset price of noise traders' demand for the asset, which is the derivative (with respect to the proportion of heuristic traders) of the multiplication of the pricing parameter and noise traders' demand for the asset. The other component is the impact on the asset price of the aggregate demand of rational and heuristic traders, which is defined as the derivative (with respect to the proportion of heuristic traders) of the multiplication of the pricing parameter and the aggregate demand of rational and heuristic traders. This impact is positive in the presence of good news. Hence, the net impact of these two components determines whether the impact on the asset price of the increased proportion of heuristic traders is positive or negative (or the derivative of the asset price with respect to the proportion of heuristic traders is positive or negative). If it is positive, then the asset price goes up as the proportion of heuristic traders increases. If rational traders is buying the asset in responding to good news, this means that the asset price is driven up higher.
than what it would be if the market consists of only rational traders along with noise traders. That is, the asset price overreacts to the good news. If rational traders is selling the asset in responding to good news, this, together with the positive derivative of the asset price with respect to the proportion of heuristic traders, implies that the asset price is pushed down not as low as it would if the market consists of only rational traders along with noise traders. That is, the asset price underreacts to the good news. On the other hand, if the derivative of the asset price with respect to the proportion of heuristic traders is negative, then when rational traders are selling the asset, the asset price is pushed lower than it would if the market consists of only rational traders along with noise traders. The asset price overreacts to good news in this case. When the derivative of the asset price with respect to the proportion of heuristic traders is negative, the asset price will underreact to good news if rational traders are buying the asset in responding to good news. Similar explanations would apply to explaining why the asset price can underreact or overreact to some bad news. The remainder of this paper consists of three sections. The next section describes the model. Section 3 presents the results. Section 4 concludes the paper.

II. The Model

Consider a one-period model of an asset market with one asset and one market maker. The market maker supplies the liquidity to the market. The cost of doing so is assumed to be zero for the simplicity. All traders submit their market orders for the asset to the market maker. It is common belief that the payoff of the asset is normally distributed with the mean of $\bar{\theta}$ and variance of $\sigma^2$. No trader knows the payoff of the asset but rational and heuristic traders receive an informational signal about the asset’s payoff before any trade occurs. This informational signal is modeled as $S = \theta + \varepsilon$ where $\varepsilon$ is normally distributed with the mean of zero and variance of $\sigma^2$. The random variables $\theta$ and $\varepsilon$ are independent. The informational signal is considered as good news if $S > \bar{\theta}$ and it is considered as bad news if $S < \bar{\theta}$.

After receiving the informational signal about the asset’s payoff, rational traders updates their conditional mean about the asset’s payoff according to the following:

$$E(\theta | (S, r)) = \bar{\theta} + \frac{\sigma^2_r}{\sigma^2 + \sigma^2} (S - \bar{\theta}),$$

where $r$ indicates rational traders.

Heuristic traders exhibit representativeness. The representativeness heuristic is a type of behavioral bias. It is well documented in the psychologists’ experiments (see Kahneman and Tversky (1973), Tversky and Kahneman (1974), and Grether (1980)). In this model, traders with representativeness place too much weight on the current information and too little weight on their prior knowledge when they update their beliefs about the asset’s payoff. Hence, heuristic traders’ updated conditional mean of the asset’s payoff is modeled as the following:

$$E(\theta | (S, h)) = \bar{\theta} + m(E_r(\theta | S) - \bar{\theta}) = \bar{\theta} + \frac{m\sigma^2_r}{\sigma^2 + \sigma^2} (S - \bar{\theta}),$$

where $h$ indicates heuristic traders and $m > 1$. The conditional mean of the asset’s payoff for heuristic traders is larger than that for rational traders when the informational signal indicates good news (i.e., $S > \bar{\theta}$); and it is smaller than that for rational traders when the informational signal indicates bad news (i.e., $S < \bar{\theta}$). Here, heuristic traders underweight their prior knowledge. This is consistent with the base-rate underweighting or base rate fallacy characterized in Kahneman and Tversky (1973), Tversky and Kahneman (1974), and Grether
Both rational and heuristic traders are considered as informed traders. It is assumed that there are $N$ informed traders. The proportion of informed traders being heuristic traders is denoted as $f$, where $f \in [0,1]$. The market maker behaves competitively. After receiving the aggregate demand of all traders, he sets the asset price equal to the expected asset's payoff conditional on the observed aggregate demand for the asset. The asset price is denoted as $P$ and the aggregate demand is denoted as $D$. Hence, the asset price is determined by the following equation:

$$P = E(\theta|D).$$

(3)

The equilibrium is characterized by the following: (1) Given the asset pricing rule stated in equation (3) and taken into account the impact of his market order on the asset price and on other traders' market orders, trader $i$, where $i \in \{1,2,\ldots,N\}$, of type $j$, $j = r, h$, chooses his market order (denoted as $X_{ij}$) to maximize his expected profit

$$\max_{X_{ij}} [E(\theta(S,j)) - E(P(S,X_{ij}))]X_{ij},$$

(4)

where $E(\theta(S,j)) = E(\theta(S,r))$ if $j = r$; $E(\theta(S,j)) = E(\theta(S,h))$ if $j = h$, and $E(P(S,X_{ij})) = E(\theta|D)$. (2) Given all the market orders coming from all traders, the market maker sets the asset price equal to the expected asset's payoff according to equation (3).

Since all traders of the same type have the same equilibrium strategy, the equilibrium market order for rational and heuristic traders are denoted as $X_r$ and $X_h$. As shown in the appendix, the equilibrium strategies for rational and heuristic traders, and the equilibrium asset price are as follows.

$$X_r = \frac{\eta(Nf-Nfm+1)(S-\bar{\theta})}{\lambda(N+1)},$$

(5)

$$X_h = \frac{\eta(N(1-f)(m-1)+m)(S-\bar{\theta})}{\lambda(N+1)},$$

(6)

and

$$P = \bar{\theta} + \lambda x + \frac{N\eta(1-f+fm)(S-\bar{\theta})}{N+1},$$

(7)

where

$$\lambda = \left(\frac{N\eta\sigma_\theta^2(fm-f+1)(Nf(1-m)+1)}{\sigma_\theta^2(1+N)^2}\right)^{\frac{1}{2}}$$

(8)

and the positive root is used to ensure the second order condition of the optimization problem holds and ensure the asset price is an increasing function of the aggregate demand of all traders.

Remark 1: Note that rational and heuristic traders are risk neutral in this model.

Remark 2: Note from equation (7) that the parameter $\lambda$ is a function of the proportion of heuristic traders.

Remark 3: Note from equations (5) and (6) that in responding to good news, rational traders' demand for the asset price can be positive or negative depending on the model parameter. However, in responding to good news, heuristic traders' demand for the asset is positive. This is because the conditional mean of the asset's payoff for heuristic is larger than that for rational traders in responding to good news due to representativeness heuristic (see equations (1) and (2)). On the other hand, heuristic traders' demand for the asset is negative in responding to bad
news since the conditional mean of the asset's payoff for heuristic traders is smaller than that for rational traders in the case of bad news (see equations (1) and (2)).

Using equations (5) and (6),

\[ X_h - X_r = \frac{(m-\eta+N\eta(m-1))(\bar{s}-\bar{\theta})}{\lambda(N+1)} \]  

(9)

Equation (9) suggests that the demand coming from heuristic traders are larger than that coming from rational traders in the case of good news and smaller than that coming from rational traders in the case of bad news. This is due to the fact that the conditional mean of the asset's payoff for heuristic is larger (smaller) than that for rational traders in case of good (bad) news due to representativeness heuristic (see equations (1) and (2)).

III. The Results

This section presents detailed analysis on how the asset price can overreact to some new information and underreact to others.

In this framework, the asset price overreaction to new information occurs when the asset price, in responding to new information, is higher (lower) than what it would be if the market consists of only rational traders (along with noise traders) \((m = 1, f = 1)\), who are buying (selling) the asset; otherwise, the asset price underreaction to new information occurs.

The occurrence of whether the asset price overreacts or underreacts to new information depends on whether the asset price increases or decreases as the proportion of heuristic traders increases. For this reason, the following two equations are computed from equations (7) and (8):

\[ \frac{d\lambda}{df} = \frac{N\eta\sigma_2^2(1-m)(N-2Nf+2Nfm-1)}{2\lambda(N+1)^2\sigma^2} \]  

(10)

and

\[ \frac{dp}{df} = N\eta(m-1)(\bar{s}-\bar{\theta}) \]  

(11)

Note from equation (10) that since \( N - 2Nf + 2Nfm - 1 > 0 \) for \( m > 1 \), it follows that \( \frac{d\lambda}{df} < 0 \).

Also, note from equation (11) that the impact on asset price of the increased proportion of heuristic traders can be positive or negative depending on the net effect of two components. One component is the impact on the asset price of noise traders' demand (which is the first term of the right hand side of equation (11). This impact is positive if noise traders are net sellers and otherwise, it is negative. The other component, which is second term of right hand side of equation (11), is the impact on the asset price of the aggregate demand of rational and heuristic traders. This impact is positive if the informational signal indicates good news and it is negative if the informational signal indicates bad news. If the net effect of these two impacts is positive (negative), then the asset price increases (decreases) as the proportion of heuristic traders increases.

The following analyzes four scenarios. These scenarios are classified according to whether noise traders are net buyers or sellers and whether informational signal indicates good news or bad news.

**Scenario (a): Informational signal indicates good news and noise traders are net sellers**

In this case, equation (11) implies that \( \frac{dp}{df} > 0 \) always holds. This is because the impact on the asset price of noise traders' negative demand and of the aggregate demand of rational and
heuristic traders are both positive. However, the demand for the asset coming from rational traders can be positive or negative. In responding to good news, rational traders buy the asset if \( N_f - Nfm + 1 > 0 \) or \( f < \frac{1}{N(m-1)} \) (see equation (5)). This, together with the positive derivative of the asset price with respect to the proportion of heuristic traders, implies that the asset price is driven up higher than what it would be if the market consists of only rational traders (who are buying the asset) along with noise traders. That is, the asset price overreacts to good news. On the other hand, if \( N_f - Nfm + 1 < 0 \) or \( f > \frac{1}{N(m-1)} \) and \( m > 1 + \frac{1}{N} \) then rational traders sell the asset in responding to good news. This, together with the positive derivative of the asset price with respect to \( f \), implies that the asset price is pushed down not as low as it would be if the market consists of only rational traders (who are selling the asset) along with noise traders. That is, the asset price underreacts to the good news.

**Scenario (b): Informational signal indicates bad news and noise traders are net buyers**

In this case, equation (11) implies that \( \frac{dp}{df} < 0 \) is always true. This is because the impact on the asset price of noise traders' positive demand and of the aggregate demand of rational and heuristic traders are both negative. On the other hand, the rational traders can be buying the asset or selling the asset depending on the model parameters values. If \( N_f - Nfm + 1 > 0 \) or \( f < \frac{1}{N(m-1)} \) equation (5) implies that rational traders sell the asset in responding to bad news. In this case, the negative derivative of the asset price with respect to \( f \) means that the asset price is pushed down lower than what it would be if the market consists of only rational traders (who are selling the asset) along with noise traders. This is the asset price overreaction to the bad news. On the other hand, if \( N_f - Nfm + 1 < 0 \) or \( f > \frac{1}{N(m-1)} \) and \( m > 1 + \frac{1}{N} \) rational traders buy the asset (see equation (5)). Hence, the negative derivative of the asset price with respect to \( f \) means that the asset price is driven up not as high as it would be if the market consists of only rational traders (who are buying the asset) along with noise traders. This is the asset price underreaction to the bad news.

**Scenario (c): Informational signal indicates good news and noise traders are net buyers**

In this case, the impact on the asset price of noise traders' positive demand is negative while the impact on the asset price of the aggregate demand of rational and heuristic traders is positive. If the informational signal is sufficiently large so that the impact on the asset price of the aggregate demand of rational and heuristic traders dominates that of noise traders' positive demand (i.e., \( S > \bar{\theta} + \frac{\bar{\sigma}^2 N_f(1-m)+1-N}{2\lambda(N+1)\bar{\sigma}^2} \)), the equation (11) implies that \( \frac{dp}{df} > 0 \). On the other hand, rational traders will buy the asset in responding to good news if \( N_f - Nfm + 1 > 0 \) (or \( f < \frac{1}{N(m-1)} \)) (due to equation (5)). This, together with the positive derivative (with respect to \( f \)) of the asset price, implies that the asset price is driven up higher than what it would be if the market consists of only rational traders (who are buying the asset) along with noise traders. That is, the asset price overreacts to the sufficiently good news. However, rational traders will sell the asset in responding to good news if \( N_f - Nfm + 1 < 0 \) (or \( f > \frac{1}{N(m-1)} \)) and \( m > 1 + \frac{1}{N} \) (see equation (5)). Together with \( \frac{dp}{df} > 0 \), this implies that the asset price is pushed down not as low as it would if the market consists of only rational traders along with noise traders. This is the asset price
underreaction to the sufficiently good news.

Now, if the informational signal is sufficiently small so that the impact on the asset price of the aggregate demand of rational and heuristic traders is dominated that of noise traders' positive demand (i.e., $\bar{\theta} < S < \bar{\theta} + \frac{\sigma^2 \alpha}{2 \lambda (N+1) \sigma^2}$), then equation (11) implies that $\frac{dp}{df} < 0$. Since rational traders buy the asset in responding to good news if $Nf - Nfm + 1 > 0$ (or $f < \frac{1}{N(m-1)}$) (see equation (5)), this, together with the negative derivative of the asset price respect to $f$, implies that the asset price is driven up not as high as it would be if the market consists of only rational traders along with noise traders. This is the asset price overreaction to the mildly good news. On the other hand, the rational traders sell the asset in responding to good news if $Nf - Nfm + 1 < 0$ (or $f > \frac{1}{N(m-1)}$ and $m > 1 + \frac{1}{N}$) (see equation (5)). Then, the asset price is pushed down lower than what it would be if the market consists of only rational traders (who are buying the asset) along with noise traders. This is the asset price underreaction to the mildly good news.

**Scenario (d): Informational signal indicates bad news and noise traders are net sellers**

In responding to bad news, the impact on the asset price of the aggregate demand of rational and heuristic traders is negative while the impact on the asset price of noise traders' negative demand is positive. If the informational signal is small negative number so that the impact on the asset price of noise traders' negative demand dominates that of the aggregate demand of rational and heuristic traders (i.e., $\bar{\theta} > S > \bar{\theta} + \frac{\sigma^2 \alpha}{2 \lambda (N+1) \sigma^2}$), then equation (11) implies that $\frac{dp}{df} > 0$. Furthermore, rational traders buy the asset in responding to bad news if $Nf - Nfm + 1 < 0$ (or $f > \frac{1}{N(m-1)}$ and $m > 1 + \frac{1}{N}$) (see equation (5)). Hence, the asset price is driven up higher than what it would be if the market consists of only rational traders (who are buying the asset) along with noise traders. This is the asset price overreaction to the mildly bad news. However, in responding to bad news, rational traders will sell the asset if $Nf - Nfm + 1 > 0$ (or $f < \frac{1}{N(m-1)}$) (see equation (5)). This, together with positive derivative (with respect to $f$) of the asset price implies that the asset price is pushed down not as low as it would be if the market consists of only rational traders (who are selling the asset) along with noise traders. This is the asset price underreaction to the mildly bad news.

Now, if the informational signal is large negative number so that the impact on the asset price of noise traders' negative demand is dominated that of the aggregate demand of rational and heuristic traders (i.e., $S < \bar{\theta} + \frac{\sigma^2 \alpha}{2 \lambda (N+1) \sigma^2}$), then equation (11) implies that $\frac{dp}{df} < 0$. Since rational traders buy the asset in responding to bad news if $Nf - Nfm + 1 < 0$ (or $f > \frac{1}{N(m-1)}$ and $m > 1 + \frac{1}{N}$) (see equation (5)), this together with the negative derivative of the asset price with respect to $f$, implies that the asset price is driven up not as high as it would be if the market consists of only rational traders (who are buying the asset) along with noise traders. This is the asset price overreaction to the bad news. However, in responding to bad news, rational traders will sell the asset if $Nf - Nfm + 1 > 0$ (or $f < \frac{1}{N(m-1)}$) (see equation (5)). This together with the negative derivative of the asset price with respect to $f$, implies that the asset price is pushed down lower than what it would be if the market consists of only rational traders (who are selling the asset) along with noise traders. This is the asset price overreaction to the sufficiently good news.
bad news.

The results from the above analysis are summarized in the following two propositions. Proposition 1 presents the results associated with how the asset price reacts to good news and how the asset price reacts to bad news is presented in Proposition 2.

**Proposition 1**
(1) When noise traders are net sellers, if \( f < \frac{1}{N(m-1)} \), the asset price overreacts to any good news; if \( f > \frac{1}{N(m-1)} \) and \( m > 1 + \frac{1}{N'} \), the asset price underreacts to any good news. (2) When noise traders are net buyers and the informational signal indicates good news with its realization \( S > \bar{\theta} + \frac{x \sigma_q^2 (2Nf(1-m)+1-N)}{2\lambda(N+1)\sigma_x^2} \), if \( f < \frac{1}{N(m-1)} \), the asset price overreacts to this good news; if \( f > \frac{1}{N(m-1)} \) and \( m > 1 + \frac{1}{N} \), the asset price underreacts to this good news. (3) When noise traders are net buyers and the informational signal indicates good news with its realization \( \bar{\theta} > S < \bar{\theta} + \frac{x \sigma_q^2 (2Nf(1-m)+1-N)}{2\lambda(N+1)\sigma_x^2} \), if \( f < \frac{1}{N(m-1)} \), the asset price underreacts to this good news; if \( f > \frac{1}{N(m-1)} \) and \( m > 1 + \frac{1}{N} \), the asset price overreacts to this good news.

Similar results are stated below when the informational signal suggests bad news.

**Proposition 2**
(1) When noise traders are net buyers, if \( f < \frac{1}{N(m-1)} \), the asset price overreacts to any bad news; if \( f > \frac{1}{N(m-1)} \) and \( m > 1 + \frac{1}{N'} \), the asset price underreacts to any bad news. (2) When noise traders are net sellers and the informational signal indicates bad news with its realization \( \bar{\theta} > S > \bar{\theta} + \frac{x \sigma_q^2 (2Nf(1-m)+1-N)}{2\lambda(N+1)\sigma_x^2} \), if \( f < \frac{1}{N(m-1)} \), the asset price underreacts to this bad news; if \( f > \frac{1}{N(m-1)} \) and \( m > 1 + \frac{1}{N} \), the asset price overreacts to this bad news. (3) When noise traders are net sellers and the informational signal indicates bad news with its realization \( S < \bar{\theta} + \frac{x \sigma_q^2 (2Nf(1-m)+1-N)}{2\lambda(N+1)\sigma_x^2} \), if \( f < \frac{1}{N(m-1)} \), the asset price overreacts to this bad news; if \( f > \frac{1}{N(m-1)} \) and \( m > 1 + \frac{1}{N} \), the asset price underreacts to this bad news.

The results of both propositions suggest that representativeness heuristic is capable of generating asset price underreaction to both good news and bad news in addition to asset price overreaction to both good news and bad news.

The following section presents some concluding remarks.

**IV. Conclusion**
In this simple one-period model with one asset and one market maker. The payoff of the asset is unknown to all market participants. But rational and heuristic traders receive an informational signal about the asset's payoff before any trade takes place. The market maker supplying the liquidity to the market. Traders are risk neutral and they act strategically to maximize their expected profits. In this model, traders submit the market orders to the market maker. After observing the aggregate demand of all traders (including noise traders' demand for the asset), the market maker competitively sets the asset price equal to the expected asset's payoff conditional on the observed aggregate demand of all traders. The equilibrium notion in this market is essentially Cournot equilibrium. This paper is able to obtain a unique equilibrium where the asset
price overreacts to some informational signals with certain model parameters conditions and
underreacts to the same informational signals or different information signals with other model
parameters restrictions.

It is worth emphasizing that this paper does not attempt to explain the empirical anomaly of asset
price underreaction and overreaction to new information. Instead, it focuses on demonstrating
that representativeness heuristic is capable of generating asset price underreaction to new
information in an asset market with strategic interaction among traders.

References
Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment,
Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and
Douks, John and Phillip McKnight, 2005, European Momentum Strategies, Information
Accounting and Economics, 27(1), 89-124.
Quarterly Journal of Economics 95, 537-557.
Hirshleifer, David, Sonya Lim, and Siew Hong Teoh, Limited Investor Attention and Stock
Market Misreaction to Accounting Information, forthcoming, Review of Asset Pricing
Studies.
Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum
Jegadeesh, Narasimhan and Sheridan Titman, 2001, Profitability of Momentum Strategies: An
Kahneman, D., and A. Tversky, 1973, "On the psychology of prediction", Psychological Review
80, 237-251.
Luo, G. 2011, "Can Heuristic Traders Survive in a Competitive Securities Market?", Working
Paper, McMaster University.
Biosciences, 40, 145-156.
Science 185, 1124-1131.

Appendix
Derivation of equations (1) and (2): Notice that $S = \theta + \epsilon$, where $\theta$ is normally distributed
with mean $\bar{\theta}$ and standard deviation of $\sigma_\theta$ and $\epsilon$ is also normally distributed with mean zero and
standard deviation of $\sigma_\epsilon$; furthermore, $\theta$ and $\epsilon$ are independent. Hence, the following are true, (a)
$S$ and $\theta$ are jointly normal distributed; (b) $Var(S) = \sigma_\theta^2 + \sigma_\epsilon^2$; (c) $Cov(\theta,S) = \sigma_\theta^2$. Result (c)
comes from the following $Cov(\theta,S) = E[(\theta - \bar{\theta})(S - \bar{S})] = E(\theta S) - \bar{\theta}^2 = E(\theta(\theta + \epsilon)) -
\bar{\theta}^2 = E\theta^2 - \bar{\theta}^2 = \sigma_\theta^2$. With the results (a), (b) and (c), equation (1) and (2) follows from a
proposition stated below. If the random variables $X^*$ and $Y^*$ are jointly normally distributed, then

$$E(X^*|Y^* = Y) = EX^* + \frac{\text{Cov}(X^*, Y^*)}{\text{Var}(Y^*)}(Y - EY^*)$$

and

$$\text{Var}(X^*|Y^* = Y) = \text{Var}(X^*) - \frac{[\text{Cov}(X^*, Y^*)]^2}{\text{Var}(Y^*)}$$

(See Hoel, p.200).

**Derivation of equations (5), (6) and (7):** Denote the total number of heuristic traders as $N_h$ and the total number of rational traders as $N_r$. Hence, $N = N_r + N_h$. Assume that the equilibrium strategies for rational and heuristic traders are linear functions of their informational signal and they are as follows:

$$X_{ir} = a_{ir} + b_{ir} S,$$

for $i = 1, 2, \ldots, N_r$. And

$$X_{ih} = a_{ih} + b_{ih} S,$$

for $i = 1, 2, \ldots, N_h$. Also assume that the equilibrium asset price follows the linear pricing rule:

$$P = \mu + \lambda D,$$

where $D = \sum_{i=1}^{N_r} X_{ir} + \sum_{i=1}^{N_h} X_{ih}$; and all the coefficients $\mu$, $\lambda$, $a_{ir}$ and $b_{ir}$ for $i = 1, 2, \ldots, N_r$; $a_{rh}$ and $b_{ih}$ (for $i = 1, 2, \ldots, N_h$) are to be determined later.

Substituting equations (1), (2), (12), 13 and (14) into the optimization problem (4), it follows that the first order condition for the optimization problem (4) is as follows: for

$$\eta = \frac{\sigma^2}{\sigma^2 + \sigma^2},$$

$$\bar{\theta} + \eta(S - \bar{\theta}) - \mu - \lambda(2X_{ir} + \sum_{n=1}^{N_r} (a_{nr} + b_{nr} S) + \sum_{n=1}^{N_h} (a_{nh} + b_{nh} S)) = 0,$$  \hspace{1cm} (15)

and

$$\bar{\theta} + m\eta(S - \bar{\theta}) - \mu - \lambda(2X_{ih} + \sum_{n=1}^{N_h} (a_{nh} + b_{nh} S) + \sum_{n=1}^{N_r} (a_{nr} + b_{nr} S)) = 0.$$ \hspace{1cm} (16)

Again, substituting equation (12) and (13) into equation (15) and (16) respectively, it follows that

$$a_{ij} = \frac{\bar{\theta} - \mu - R_{ij} \eta \bar{\theta}}{\lambda} - A,$$  \hspace{1cm} (17)

and

$$b_{ij} = \frac{R_{ij} \eta}{\lambda} - B,$$  \hspace{1cm} (18)

where $A = \sum_{n=1}^{N_r} a_{nr} + \sum_{n=1}^{N_h} a_{nh}$, and $B = \sum_{n=1}^{N_r} b_{nr} + \sum_{n=1}^{N_h} b_{nh}$.

Notice from equations (17) and (18), that for $i' \neq i$, $a_{ij} = a_{i'j}$ and $b_{ij} = b_{i'j}$ for the same $j \in \{r, h\}$ (the same type of traders). Hence, let $a_{ir} = a_r$, $b_{ir} = b_r$ when $R_j = 1$; and $a_{ih} = a_h$, $b_{ih} = b_h$ when $R_j = m$. Equations (17) and (18) imply the following four equations are true:

$$a_r = \frac{\bar{\theta} - \mu + \eta \bar{\theta}(N_f(m-1)-1)}{\lambda(N+1)},$$ \hspace{1cm} (19)

$$a_h = \frac{\bar{\theta} - \mu + \eta \bar{\theta}(N(1-f)(1-m)-m)}{\lambda(N+1)},$$ \hspace{1cm} (20)

$$b_r = \frac{\eta(1+NF(1-m))}{\lambda(N+1)},$$ \hspace{1cm} (21)

and

$$b_h = \frac{\eta(m-N(1-f)(1-m))}{\lambda(N+1)}.$$ \hspace{1cm} (22)
Using equation (3),
\[ P = E(\theta|A + BS + x = D) \]
\[ = \bar{\theta} + \frac{B\sigma_0^2}{B^2\sigma_0^2+\sigma_x^2}(D - A - B\bar{\theta}) \].

Using equations (14) and (23) along with the definitions of A and B, one can show that
\[ \mu = \bar{\theta} \],

and
\[ \lambda^2 = \frac{N\mu\sigma_0^2(f-m+1)(Nf(1-m)+1)}{\sigma_x^2(1+N)^2} \].

Therefore, using equations (19) through (22), (24), (12) and (13), the equilibrium strategies for rational and heuristic traders and the equilibrium asset price for the market maker follows. Note that \( \lambda \) is determined by equation (25) and the positive root from equation (25) is used to ensure that the second order condition of the optimization problem (4) holds and ensure that the equilibrium price is increasing in the total demand for the asset.
Data Visualization: An Alternative and Complementary Learning Strategy to Teaching Ratio Analysis

Dr. David DeBoskey
KPMG Faculty Fellow
The Charles W. Lamden School of Accountancy
College of Business Administration
5500 Campanile Drive, San Diego, CA 92181-8221
debocpafa2005@aol.com

Martha Doran, Ph.D.
Associate Professor
San Diego State University
Abstract

Financial statement analysis depends on a large degree to using a comparison strategy such as ratio analysis. Ratio analysis is a fundamental topic taught in most introductory financial accounting classes (at both the undergraduate and graduate level). We offer an alternative approach for evaluating financial performance of firms across multiple financial dimensions utilizing radar graphs (a/k/a spider graphs) a common graphing mechanism found in Microsoft Excel. We also present a tutorial on radar graphs that offers a step-by-step approach to deploying them. Specifically, our research explores whether the influence of information presentation affects the perceived and actual learning outcomes of our subjects. Our students’ self-reported assessments demonstrate that they clearly viewed radar graphs as a complementary tool that enhanced their overall understanding of firm performance across a wide-array of measures. The independent learning assessments demonstrate that students understand radar graphs and can effectively use them to analyse and summarize firm performance across multiple dimensions. Overall, we provide some evidence that radar graphs integrate and present profitability, liquidity, solvency, risk/leverage, efficiency and cash flow in a manner recognizable to students that enhance students’ ability to correctly summarize and interpret firm performance.

Keyword: ratio analysis, data visualization, financial statement analysis, classroom assessments

I. Introduction: Motivation for the use of data visualization

The American Institute of Certified Public Accountants (AICPA’s) expressed a need for accountants to demonstrate an ability to see the “big picture” of a company’s operations, as one of the core competencies in its 2009 publication. Research has shown that less experienced users’ process financial information in an unstructured manner, demonstrating a lack of understanding among the various financial statement items (Fredrickson & Miller 2004, Hunton & McEwen, 1997). Accounting researchers have long anticipated financial information tools that help users develop more comprehensive understanding of how changes in various financial items affected a company’s overall operations (Wallman, 1997). However, such tools have been given little research attention. We plan to respond to this significant opportunity by demonstrating that data visualization tools are helpful in teaching the substance and interaction of important accounting concepts to students.

Importance to Accounting Education

The extant accounting education research has been sparse on the use of data visualization techniques to teach ratio analysis. Our personal classroom experience shows students are often able to recite specific ratios by using tables and charts but have limited ability to interpret what the ratios could indicate, especially when multiple ratios are being analyzed. Hill & Milner, (2003) argue the importance of integrating graphical financial reporting into accounting education. Graphical visualization has been shown to improve the efficiency and effectiveness of data analysis (Volmer, 1992).

Moreover, recent empirical archival research has explored the voluntary disclosure practices contained in firms’ annual reports (Dilla & Janvrin, 2010). They find that graph use is more (less) pronounced when the magnitude and direction of firm performance is positive (negative). Therefore, firms’ graph use is less when performance trends are not as favorable. It would appear that firm graph use enables readers to more readily ascertain both positive and negative trends, and that firms tend to avoid graphing negative performance trends.

The adage of ‘a picture being worth a 1000 words’ motivates the value of radar graphs as a tool that may help students see what they are trying to explain and in turn gain confidence in their
interpretation of the financial analysis. At a minimum, using radar graphs to give a pictorial representation of ratio results, allows students to better frame questions, both about their understanding of what the ratios are saying about the company’s performance and about the performance issues potentially underlying the ratios, which is one of the key components of value provided by ratio analysis.

Based on a comprehensive review of the accounting education research we were unable to identify specific research paper(s) that utilized electronic spreadsheet capabilities such as those found in Microsoft Excel to complement and simplify the process of performing a detailed financial review of firms’ consolidated financial statements. Our research extends the extant accounting education research by offering specific guidance and tools for the implementation of radar graphs in teaching financial ratio analysis and has four specific contributions: (i) to offer a generalized approach for embedding radar graphs into teaching graduate students ratio analysis; (ii) to offer an overview of radar graphs and their practicality in teaching ratio analysis including a scenario for comparing a companies’ results across several important ratio groupings: liquidity, profitability, solvency, efficiency and cash flows (iii) to highlight the potential incremental benefits of radar graphs over complex tabular numerical presentations of firm performance; and (iv) to provide student assessment outcomes for both regular and executive M.B.A. student subjects that offer support for this alternative and complementary learning strategy to teaching basic ratio analysis.

In the next section of the paper, we provide motivation for and an overview of data visualization techniques and the methods deployed to implement radar graphs in the classroom. The third section provides an overview of the survey instrument used to perform our student assessments. The fourth section presents further reflection on the process and results. In the final section, we discuss implications of our research as well as alternative applications to other judgment and decision-making audit task settings such as: audit risk assessment and analytical review procedures.

**Motivation for data visualization techniques**

There is a clear stream of empirical research that demonstrate the visual sense is a dominant human sense and that financial information in graphs, rather than numerical tables, may be easier for annual report users to process (See for example, DeSanctis & Jarvenpaa, 1989; Ackerman, 1991; Vessey & Galletta, 1991; Ware, 2004.). Moreover, research has shown that spatial intelligence enhances user decisions over linguistic intelligence alone (Ackerman, 1991; Tufte, 1997). Therefore, a challenging task for many educators is to assure that students comprehend vast amounts of financial information found in company annual reports in the most efficient manner. We in the education field strive not only to impart technical skills such as quantitative analyses to our students but also the ability to interpret and apply the analyses in their current and future professional careers. Seasoned accountants and professional/sophisticated investors undoubtedly are able to extract relevant details from a set of consolidated financial statements. However, more inexperienced/unsophisticated users’ like the many students we face in our graduate classrooms may lack the ability and training to perform many of the complex and integrated tasks required to properly analyze financial statements in a meaningful and persuasive manner. The ability to perform in today’s competitive and nimble workforce requires professionals to be able to reduce large amounts of financial information into a meaningful and persuasive story about a firm and its industry. This requires educators to offer students additional
tools to perform these data reduction and summary tasks. The visualization process, when used as a complementary tool to help unlock the financial ratios, can provide students with the ability to connect the dots of a company’s business strategy with the results of the business transactions. This technique can help bring the company to life for students, so they are better able to frame questions about the meaning of the ratios and about the implications to the company of the ratios. Data visualization techniques have been shown to influence nonprofessional investors’ decision-making abilities (e.g., Dilla, Janvrin & Jeffrey, 2009). Such a visualization tool—radar graphs, currently exist in Microsoft Excel. Radar graphs offer the added benefits of data reduction and efficiency. Please refer to Appendix A for a detailed description and step-by-step example on creating radar graphs.

Radar graphs can be used to summarize firm performance across multiple dimensions, such as: profitability, liquidity, solvency, efficiency and cash flows. The general approach to ratio analysis is to plug numbers into common formulas and then summarize them in a tabular numerical format, as follows:

Figure 1 presents a traditional ratio analysis that includes numbers (outputs from ratios) summarized in numerical/tabular format and for comparison purposes a radar graph that summarizes the same financial information.

By themselves, the numerical outputs from the ratios are not overly useful. The financial performance of firms’ and industries are best summarized cross-sectionally in a time-series. Although tabular presentations can accomplish this type of analysis it requires an inordinate amount of knowledge of the technical aspects of consolidated financial statements. A benefit of data visualization techniques, such as radar graphs, is that it allows students to utilize spatial intelligence to complement their linguistic intelligence.

Implementing Radar Graphs: The Basic Mechanics of Creating Them

First, we analyze the companies' financial statements from 6 points of view including: profitability, liquidity, solvency, efficiency and cash flows. Next we assess each company on the factors, and rank the companies. Then we graph each company’s ranks in the radar chart. If the company is ranked 1st among the 6 companies, we plot on the outside area of the radar graph. On the contrary, if it is ranked low on the factor, we plot near the inside area of the radar graph.

Figure 2 presents a radar graph that has six possible rankings.

---

29 This personal observation is based on our 10+ years of experience teaching executive M.B.A. students. Executive students are especially sensitive to the practicality of their M.B.A. training and prefer methods and tools that allow them to implement such methods into their daily jobs. Failure to offer executive students such tools results in suboptimal satisfaction with executive student audiences.

30 In our study, we evaluate students’ ability to assess financial performance across multiple dimensions. See Table 6 for a cross-walk of students’ self-reported assessments and independent learning outcomes.

31 Based on our collective graduate teaching experience, we have observed that students seem to have a difficult time translating the detailed and sometimes copious quantitative analyses into a “BIG PICTURE” story about the firms and industries that they just analyzed. The student self-reported assessments and independent learning assessments [see Section 4] offer some evidence of the power of radar graphs to improve students’ interpretation and application capabilities.

32 The number of performance measures will vary from analysis to analysis. This list is not meant to be an exhaustive list of performance measures. The performance measures deployed should be those that best meet the objectives of the financial review being performed.

33 For large sample of firms an alternative radar graph approach would be to create quintiles or deciles of financial performance and then rank the firms’ according to the quintile or decile they occupy for a particular financial ratio.
As you can see from figure 2 firms that have rankings on the outside area of the radar graph are those that rank highest on that factor. Three (although there could be more as well) general performance trends possibly emerge and can be observed cross-sectionally.

Panel A represents a firm that is poor at all factors. Panel B represents a firm that is very good at all factors. Panel C represents an unbalanced firm where they are good at three factors and bad at three factors.

In ranking firms’ performance it’s important to consider the ratio itself before assigning a ranking. For example, in the solvency group a company is ranked high (near outer border) on leverage if it has little of it (i.e., low leverage) and is ranked low (near center) if it has high leverage. Alternatively, all ratios can be ranked from high to low and then interpreted accordingly. Continuing with the leverage example, if a firm was ranked high in leverage (plotted on outside area) it would imply high leverage (however it may be measured). Therefore, if a firm had a high leverage ranking (meaning high leverage) in conjunction with high profitability then one could imply that the firm’s profitability is being boosted by the use of debt/leverage. There are many possible ways to prepare an analysis utilizing radar graphs. In one instance, it is possible to create radar graphs that capture firm performance for a single firm over multiple periods (single-firm—multiple-period radar graph). In another instance, it is possible to create radar graphs that capture firm performance for multiple-firms over multiple dimensions for a single period (multiple-firm—multiple-dimension—single period radar graph). Figure 4 presents a radar graph for four large competitors in the supermarket industry for the year 2009 (i.e., multiple-firm—multiple-dimension—single-period radar graph).

It is outside the scope of this research to prescribe a particular type of radar graph to deploy. Rather, we are simply demonstrating the capability and effectiveness of radar graphs in summarizing firm performance. In the next section, we outline the procedures we follow in introducing students to radar graphs.

Students are not introduced to radar graphs until substantially all major ratio groupings are covered in class. Once all major ratios are covered the students are encouraged to assess firm performance in a persuasive manner. In one lecture, students are introduced to radar graphs by demonstrating how a radar graph is created and several radar graphs depicting fictitious firms are displayed to the students. The students are then asked to consider utilizing radar graphs in their group financial analysis projects. Each student is assigned to a group of five students that are required to choose an industry to analyze. Each industry has four firms (one U.S. and three International firms) assigned to it. Students are then required to assess the firm’s performance.
along several dimensions including, but not limited to, liquidity, profitability, solvency, efficiency and cash flow generation ability.

Implementation Guidance and Overview of PowerPoint Pitch Presentation Group Project and Learning Objectives

In order to ensure that students can employ radar graphs they are divided into groups of five. Each group is asked to prepare a 30 minute presentation that analyzes the financial performance of four firms (one US and three international firms). The group presentation requires students to analyze the firms’ liquidity, solvency, efficiency, profitability and cash flows. Students are asked to provide a persuasive analysis all firms’ results of operations, statement of position and cash flows for the most recent two year period. Generally, the presentation group project is assigned at the beginning of the semester so that students can select their industries and secure all relevant financial statements. Group presentations are scheduled for the last few classes of the semester. This allows students satisfactory time to digest all course objectives and technical applications.

Learning Objectives

The specific learning objectives for the group presentation project are to develop the students’ ability to:

1. Accurately assess profitability, liquidity, solvency and efficiency for individual firms’ and the overall industry.
2. Differentiate firms’ performance between good and bad and sufficiently explain major reasons for performance discrepancies.
3. Clearly and concisely articulate key year over year performance trends for all firms’ in the industry groupings.
4. Provide meaningful discussion and analysis regarding overall firm performance compared to all other firms’ and to each other in the industry grouping.
5. Summarize firm performance in a logical and persuasive fashion.

Assessment Instrument

Two forms of assessment were used to evaluate the efficacy of radar graphs in analyzing firm performance: self-reported learning assessment (based on responses from student surveys) and independent learning assessments (based on submitted projects that were graded according to a simple assessment rubric). Assessment results are based on several sections of M.B.A. and Executive M.B.A.-level financial reporting and analysis courses for two different instructors over 2 years. Student assessments were electronically administered following the completion of the course and after grades were issued to all students. This allows for the students to provide their assessment in a more objective fashion. Asking the students to complete the survey prior to grade issuance could potentially contaminate the findings if students’ responses were aimed at pleasing the instructor. All students were informed that their responses would be kept confidential to mitigate any potential response bias limitations. The student surveys were coded and recorded by the instructor and graduate assistant to provide for maximum inter-rater reliability. The executive students are asked to complete the instrument on the last day of class after they receive their grade in the course. The international E.M.B.A. students’ completed manual surveys due to technology limitations in China. The U.S. regular M.B.A. response rate was lower (85%) versus (100%) response rate for the manual in-class surveys. Two graduate assistants independently coded each survey and any discrepancies are reconciled by a third-party in order to bolster inter-rater reliability of the test instrument. The independent learning assessments were independently graded by two instructors; and any scoring discrepancies were discussed and reconciled.
An initial pilot study was conducted on 35 former executive M.B.A. students and 20 regular M.B.A. students and based on these pilot results certain questions were tweaked and modified that seemed problematic to both student groups. The revised survey instrument was piloted again to an alternative group of 29 former executive M.B.A. and 15 regular M.B.A. students. The results of this pilot study resulted in minor modifications to the instrument. The resulting instrument was used in the current study to assess students’ self-reported learning.

**Assessment results**

Survey data from 110 masters of business administration students (53 regular M.B.A. and 57 executive M.B.A.) in 2 class years (2009-2010) were collected to provide a basis for assessing the overall impression that students’ have of radar graphs and their effectiveness as a tool in conducting ratio analysis. The 53 regular M.B.A. students were enrolled in a traditional M.B.A. program for a large public university. The 57 executive students were from two separate classes of an international executive M.B.A. program for a nationally ranked university. All students completed the assessment tool after completing the course to allow for maximum impact of the skills learned during the course.

Table 1, Panels A and B presents descriptive statistics for both the regular M.B.A. and executive M.B.A. students enrolled across three graduate level financial accounting classes during 2009-2010.

The mean (median) age of the regular M.B.A. and [executive M.B.A] student in the sample is 27.52 (25.4) and [37.03 (36.20)]. The youngest (oldest) is 21 (38) and [26 (48)] years. The mean (median) number of years of total work experience and managerial experience for the regular M.B.A. and [executive M.B.A.] is 5.44 (4.00) [12.18 (11.00)] and 3.20 (2.00) [8.50 (7.95)], respectively. These descriptive statistics are similar to recent data released by Executive MBA Council [2007] revealing an average age of 36.3 years and 12.7 years of work experience. The sample appears to conform to the executive student profile across all members of the Executive MBA Council. For regular M.B.A. students, an international applicant survey from QS TopMBA [2010], reported the percentages of applicants who are younger and have less experience have grown over the last three years in all regions of the world. They report this younger, less experienced pool in the US and Canada has increased from 48% of all applicants in 2008 to 63% of all applicants in 2010. While this is large increase, it is smaller than the increases in other world regions. For example, Europe’s pool of younger, less experienced applicants grew from 50% to 70% during the same period. Students in all areas of the world seem to realize that graduate school is a good alternative during a global economic recession.

Table 2—Table 5 below summarizes the student assessments for both regular and executive M.B.A. subjects across three perceived measurement scales: pre-course knowledge, effectiveness, usefulness and knowledge/impact. Table 6 presents the results of the student learning outcomes assessment utilizing an expectations rubric derived from the learning objectives for the case project. Table 2 presents the student assessments regarding students’ pre-course familiarity with radar graphs. Table 3 presents the students’ perceived effectiveness of using radar graphs in ratio analysis. Table 4 presents the students’ perceived usefulness of using radar graphs in ratio analysis. Table 5 presents students’ self-assessment of the knowledge impact they gained from learning about and working with radar graphs. Table 6 presents the independent learning assessments (based on submitted case projects evaluated using a simple expectations rubric).

Table 2 reports students’ overall pre and post-course knowledge of radar graphs. Students’ are
posed with a series of statements derived from the underlying learning objectives for the alternative learning strategy; allowable responses ranged from 1 (Yes) to 2 (No) for statements one and two; and 1 (strongly disagree) to 5 (strongly agree) for statements three through five. Student responses to the statements indicate that a large percentage of students in both regular and executive M.B.A. had no prior awareness of radar graphs. For questions one and two approximately 64% (mean response of 1.64) of the regular M.B.A. and 69% (mean response of 1.69) of the executive M.B.A. had no prior exposure to radar graphs. Approximately 80 percent of all students either agree or strongly agree that they are very likely to use radar graphs in future graduate M.B.A./M.S. course work. Executive students’ agreed significantly more (0.35 on a five-point Likert scale) than regular M.B.A.s that the course and group project greatly enhanced their understanding of radar graphs. Overall, approximately 90 percent of students either agree or strongly agree that the course and group project greatly enhanced their understanding of the capabilities of radar graphs in performing ratio analysis. We argue that the executive students’ greater experience allows them to grasp new tools in more efficient manner than traditional M.B.A. students who have less overall work experience. In our sample, the mean work experience of regular M.B.A.s in years is 5.44 years vs 12.2 years for the executive students. We believe this contributes to the weaker agreement seen in the regular M.B.A. students.

Table 3 reports students’ assessment of perceived radar graph effectiveness. Students are posed with a series of statements intended to ascertain whether numerical ratios are more effective than radar graphs in understanding overall firm performance. Allowable responses were: 1 (strongly disagree) to 5 (strongly agree) for statements one through four. One item (response item 3) show that approximately 90% of students in both regular and executive M.B.A. either agree or strongly agree that the most effective way to assess firm performance is to combine numerical ratios and radar graphs. One statement (response item 1) approximately 60 percent of students’ either disagree or strongly disagree that numerical ratios are effective by themselves. However, only 40% of students’ disagree or strongly disagree that radar graphs are effective by themselves. The executive students’ responses were significantly higher (p < 0.01) in three of the four statements (response items 2, 3 and 4). The significantly higher agreement of 1.02, 0.38 and 0.77 on the five-point scale demonstrate that executives have a more favorable view radar graphs over numerical ratios. We note, however, that approximately 85 percent of students either agree or strongly agree that radar graphs are the most effective way to understand overall firm performance. Overall, the results strongly support the conclusion that radar graphs are more effective in analyzing firm performance than numerical ratios alone for both student groups.

Table 4 reports students’ assessment of perceived radar graph usefulness. Students’ are posed with a series of statements intended to ascertain whether numerical tabular ratios are more useful than radar graphs in assessing firm-specific measures such as: profitability, liquidity, efficiency and solvency trends over several reporting periods. Allowable responses were: 1 (strongly disagree) to 5 (strongly agree) for statements one through three. One item (response item 3) show that approximately 84% of students agree or strongly agree that the radar graphs are the most useful way to assess key profitability, liquidity, efficiency and solvency trends over several years. One statement (response item 1) was marginally significantly different (p < 0.10) between regular and executive M.B.A. students indicating that 58% either disagree, strongly disagree or are neutral on the use of radar graphs by themselves to assess key financial ratios over several years. It is worth noting that approximately 58% also disagree, strongly disagree or are neutral on the use of numerical ratios by themselves. However approximately 84 percent agree or strongly agree that combing numerical ratios with radar graphs are the most useful way to assess
key business ratios over several years. The executive students’ responses were significantly higher (p < 0.001) in two of the three statements (response items 2 and 3) indicating that executive students’ higher agreement (0.98 and 0.63) on the five-point scale strongly indicate the executive students’ favorability of the usefulness of radar graphs combined with numerical tabulations. The increased level of work experience of the executive students may explain this difference. Executive students are aware of business process and interactions of business events and transactions. For example, these experienced students may have participated in sales campaigns where new customers were added that did not meet existing credit standards. So their analysis of accounts receivable turn ratios, complemented by radar graphs, may build their confidence in being able to interpret the ratios, as the graphs illustrated or brought to life their experiences now shown in the ratios and graphs.

Table 5 reports students’ assessment of the knowledge impact gained from using radar graphs for assessing firm performance. Students’ are posed with a series of statements intended to measure their level of agreement with the overall level of learning obtained from using radar graphs in their group project presentations. Allowable responses were; 1 (strongly disagree) to 5 (strongly agree) for statements one through four. All four items (response items 1, 2, 3 and 4) were significantly different (p < 0.01). Approximately 75 percent of respondents either agree or strongly agree that radar graphs have made learning ratio analysis easier and that picture rather than numbers are the main reason. Approximately 76 percent of respondents also either agree or strongly agree that radar graphs had a greater impact on them during group presentations than the individual tabulated numbers. More impressive yet is that approximately 85 percent of students found that radar graphs made it easier to segregate and identify the good firms form the bad firms in their group ratio presentations. This is a major finding of our research and it points to the clear pedagogical advantage of deploying data visualization tools in ratio analysis. Finally, approximately 90 percent of students felt radar graphs enhanced the efficiency in which they spotted key trends in others’ presentations than numerical ratios alone. Overall, the knowledge impact of radar graphs seems significant across both student groups. Although, executive students’ assessment scores were significantly higher (0.60, 0.46, 0.50 and 0.63) on the five-point likert scale and indicate the greater knowledge impact gained from this student group.

Differences between Executive MBA and Regular MBA Students
It’s outside the scope of the current research design to provide copious theories and empirical support for the differences we observed between executive and regular MBA students. However, we have reported such differences for reader clarification. The main thrust of our research design is to simply examine the impact that radar graphs have on students’ perceived and actual learning outcomes. We did provide a few open-ended questions to the student respondents and we briefly discuss what they revealed below.

In analyzing a few open-ended questions from the survey instrument it appears that the executive students saw more of an opportunity to use radar graphs in their current and future jobs. Given the large percentage of full-time students of approximately 58 percent in the regular M.B.A. group helps explain why they might not agree as much as the executive students.

Summary and Conclusions
In this paper, we propose an alternative graduate teaching learning strategy for analyzing firm performance utilizing radar graphs. An overall approach to employing radar graphs is discussed and specific implementation guidance is provided. Student outcome assessments are provided for both regular and executive M.B.A. students and the results clearly show that perceived beliefs on
the impact of radar graphs is significantly altered in self-reported pre- and post-assessment surveys. Several key differences emerge between regular and executive M.B.A. students possibly indicating that prior work experience may be a moderating factor that influences the strength of these perceived beliefs. This approach offers a complementary tool to use when teaching graduate students’ ratio analysis that seems to allow students to more fully understand key performance trends across several important ratio groups.

**Limitations of the study**

This study is subject to several limitations that are typical with this stream of research including: (i) the inability to prove cause and effect regarding the effects of radar graphs on perceived student belief as well as on the added proficiency in performing a detailed firm financial analysis; (ii) studies such as these can only show association between the alternative learning strategy and students’ perceived beliefs on the added effectiveness of using radar graphs; and (iii) students’ may react favorably to any method that is perceived to make their life easier and therefore student convenience biases may obfuscate the interpretations and implications underlying this study although our large statistically significant sample potentially mitigates this potential bias. Finally, we rely primarily on student opinions on the effectiveness of radar graphs. A priori we do not measure students’ depth of their understanding about financial analysis.

**Future Research**

Future research in this area may explore the efficacy of radar graphs over traditional numerical ratios. At this juncture, little is known about how radar graphs, or visual graphs in general aid decision-makers’ overall evaluation of firm performance, including whether the speed and accuracy of such evaluations are impacted more by including complementary tools such as radar graphs. We plan to explore the use of a more direct assessment of the effect of radar graphs such as on the amount of time it takes students’ to formulate opinions with graphs versus tables and text. Additionally, we plan to explore and apply direct tests on whether radar graphs have applicability in other decision-making tasks such as audit risk assessment and analytical review procedures.
Appendix A: Creating and Radar Graph Tutorial

Step #1: Create Data Matrix (either absolute rankings or percentiles):
By Rankings (6 Firms Compared in the Cosmetics Industry):

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Revenue</th>
<th>Net Income</th>
<th>Profit from Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>5*</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2007</td>
<td>6*</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

*1-ranks lowest on that dimension; 6-ranks highest on that dimension

Step #2: Select the data in the table above (two years data)

Step #3: Insert other graphs or charts—RADAR GRAPH

Step #4: Adjust a little bit—Notice the numbers on the axis (6, 5.5, 5, 4.5 etc), click on it and choose “Set Axis Format”, then set minimum to ‘0’ and maximum to ‘6’.

Step #5: The graph should look like the following:

![Radar Graph Example]

Loreal
**Figure 1:** Common-Sized Financial Statements (Numerical/Tabular Presentation)

**Panel A: Numerical Ratios**

<table>
<thead>
<tr>
<th>Profitability</th>
<th>Industry Composite (5 major competitors)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Net Margin</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.10</td>
</tr>
<tr>
<td>2006</td>
<td>0.09</td>
</tr>
<tr>
<td><strong>Operating Margin</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.18</td>
</tr>
<tr>
<td>2006</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Gross Margin</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.62</td>
</tr>
<tr>
<td>2006</td>
<td>0.62</td>
</tr>
<tr>
<td><strong>ROE</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.12</td>
</tr>
<tr>
<td>2006</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>ROA</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.05</td>
</tr>
<tr>
<td>2006</td>
<td>0.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leverage</th>
<th>Industry Composite (5 major competitors)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D/E Ratio</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>1.94</td>
</tr>
<tr>
<td>2006</td>
<td>1.76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Solvency</th>
<th>Industry Composite (5 major competitors)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D/A Ratio</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.55</td>
</tr>
<tr>
<td>2006</td>
<td>0.55</td>
</tr>
<tr>
<td><strong>L.T. Debt/Assets</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.40</td>
</tr>
<tr>
<td>2006</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>L.T. Debt/S.E.</strong></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.64</td>
</tr>
<tr>
<td>2006</td>
<td>0.50</td>
</tr>
</tbody>
</table>
Figure 2: Radar Graphs: Alternative Performance Depictions
**Figure 3:** Radar Graphs: Alternative Performance Depictions

Panel A: Bad Company (Poor at all the factors)

Panel B: Very Good Company (Good at all the factors)

Panel C: Unbalanced Company (Good at 3 factors and bad at 3 factors)
Figure 4: Multiple-Firms—Multiple Dimensions Radar Graphs (Supermarkets, 2009)
### Table 1: Descriptive Statistics Regular and Executive MBA Students (2010)

#### Panel A: Regular MBA Students (2010)

<table>
<thead>
<tr>
<th>Executive Demographics N=53</th>
<th>INDUS TRY %</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>27.52</td>
<td>36.50</td>
<td>5.47</td>
<td>21 through 38</td>
<td></td>
</tr>
<tr>
<td>WORK EXPERIENCE (In Years)</td>
<td>5.44</td>
<td>4.00</td>
<td>4.74</td>
<td>0 through 22</td>
<td></td>
</tr>
<tr>
<td>MANAGEMENT EXPERIENCE (In Years)</td>
<td>3.20</td>
<td>2.00</td>
<td>3.32</td>
<td>1 through 6</td>
<td></td>
</tr>
<tr>
<td>ADVANCED DEGREES (%)</td>
<td>0.00</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>INDUSTRY COMPOSITION:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-Time Student</td>
<td>57.7</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Engineering (%)</td>
<td>3.80</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Finance (%)</td>
<td>3.80</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Operations (%)</td>
<td>5.80</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Sales/Marketing (%)</td>
<td>9.60</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Consulting (%)</td>
<td>1.90</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Human Resources (%)</td>
<td>3.80</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Information Technology (%)</td>
<td>7.70</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Other Technology (%)</td>
<td>5.90</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

The table presents the demographic profile of all regular MBA students including the industry composition across all regular MBA students.
### Panel B: Executive MBA Students (2010) (two sections)

<table>
<thead>
<tr>
<th>Executive Demographics</th>
<th>INDSTRY %</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>--</td>
<td>37.03</td>
<td>36.50</td>
<td>5.47</td>
<td>26 through 48</td>
</tr>
<tr>
<td>WORK EXPERIENCE (In Years)</td>
<td>--</td>
<td>12.18</td>
<td>11.00</td>
<td>3.91</td>
<td>7 through 23</td>
</tr>
<tr>
<td>MANAGEMENT EXPERIENCE (In Years)</td>
<td>--</td>
<td>7.95</td>
<td>7.30</td>
<td>3.32</td>
<td>2 through 19</td>
</tr>
<tr>
<td>PRIOR ACCOUNTING/FINANCE (# courses)</td>
<td>--</td>
<td>1.00</td>
<td>1.00</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>ADVANCED DEGREES (%)</td>
<td>28.70</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>INDUSTRY COMPOSITION: Full-Time Student</td>
<td>0.00</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Engineering (%)</td>
<td>29.90</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Finance (%)</td>
<td>18.60</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Operations (%)</td>
<td>8.50</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Sales/Marketing (%)</td>
<td>14.50</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Consulting (%)</td>
<td>2.60</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Human Resources (%)</td>
<td>5.70</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Information Technology (%)</td>
<td>10.20</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Other Technology (%)</td>
<td>10.00</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

The table presents the demographic profile of all executive students including the industry composition across all executive students.
Table 2: Classroom Assessment Data
Student’s Pre- and Post-Course Self-Reported Knowledge of Radar Graphs (n=110)

<table>
<thead>
<tr>
<th>Response Item:</th>
<th>Regular MBA Mean (Stddev) (n=53)</th>
<th>Executive MBA Mean (Stddev) (n=57)</th>
<th>Combined MBA Mean (Stddev) (n=110)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prior to enrolling in this course, I had heard of radar graphs?</td>
<td>1.64 (0.486)</td>
<td>1.69 (0.466)</td>
<td>1.67 (0.171)</td>
</tr>
<tr>
<td>2. Prior to enrolling in this course I had used radar graphs to assess firm financial performance?</td>
<td>1.90 (0.298)</td>
<td>1.81 (0.398)</td>
<td>1.86 (0.345)</td>
</tr>
<tr>
<td>3. Prior to enrolling in this course I had a strong understanding of radar graphs and their capability.</td>
<td>2.23c (1.339)</td>
<td>2.47c (1.353)</td>
<td>2.35 (1.312)</td>
</tr>
<tr>
<td>4. After completing this course and participating in the group project, I had a strong understanding of radar graphs and their capability.</td>
<td>4.25c (1.036)</td>
<td>4.60c (0.631)</td>
<td>4.42 (0.859)</td>
</tr>
<tr>
<td>5. After completing this course and participating in the group project, I am very likely to use radar graphs in my future MBA/MS courses.</td>
<td>4.22 (0.879)</td>
<td>4.37 (0.848)</td>
<td>4.31 (0.879)</td>
</tr>
</tbody>
</table>

a  A two Likert point scale was used for items one and two. The endpoints used were 1= Yes and 2= No. A five-point Likert scale was used for items three to five, the following endpoints were used: 1= strongly disagree, 5= strongly agree. A three-point Likert scale was used for item six, the following endpoints were used: 1=complementary; 2= substitute; 3=neither.

bStatistical significance based on paired t-tests. Because of potential violations of assumptions for the parametric t-tests used, we also applied nonparametric tests, which yielded very comparable results.

c Statistical difference based on paired t-tests demonstrate a significant (p<0.000) mean differences (2.02 and 2.13 on a five-point likert scale) between both regular M.B.A. and executive M.B.A.’s pre- and post- course understanding of radar graphs and their capabilities.
Table 3: Classroom Assessment Data (Perceived Radar Graph Effectiveness)  
Student’s Perceptions of Radar Graph Effectiveness  
(n=110)

<table>
<thead>
<tr>
<th>Response Item (For each of the effectiveness statements, rate your level of agreement):</th>
<th>Regular MBA Mean (Stddev) (n=53)</th>
<th>Executive MBA Mean (Stddev) (n=57)</th>
<th>Combined MBA Mean (Stddev) (n=110)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Numerical ratios by themselves are the most effective way to understand overall firm performance.</td>
<td>2.50 (0.852)</td>
<td>2.27 (1.012)</td>
<td>2.32 (0.957)</td>
</tr>
<tr>
<td>2. Radar graphs by themselves are the most effective way to understand overall firm performance.</td>
<td>3.02b (0.939)</td>
<td>4.04b (1.03)</td>
<td>3.56 (1.097)</td>
</tr>
<tr>
<td>3. Numerical ratios together with radar graphs are the most effective way to understand overall firm performance.</td>
<td>4.27b (0.717)</td>
<td>4.65b (0.653)</td>
<td>4.48 (0.700)</td>
</tr>
<tr>
<td>4. Radar graphs are more effective in synthesizing key performance trends compared to numerical (tabular) ratios alone.</td>
<td>3.83b (0.857)</td>
<td>4.60b (0.665)</td>
<td>4.24 (0.845)</td>
</tr>
</tbody>
</table>

* A five-point Likert scale was used for one to four, the following endpoints were used: 1= strongly disagree, 5= strongly agree.  
Statistical significance based on paired t-tests. Because of potential violations of assumptions for the parametric t-tests used, we also applied nonparametric tests, which yielded very comparable results. All mean differences have a p-values are <.001.
Table 4: Classroom Assessment Data (Radar Graph Usefulness)  
Student’s Perceptions of Radar Graph Usefulness  
(n=110)

Student Assessment Scores

<table>
<thead>
<tr>
<th>Response Item (For each of the usefulness statements, rate your level of agreement):</th>
<th>Regular MBA Mean (Stddev) (n=53)</th>
<th>Executive MBA Mean (Stddev) (n=57)</th>
<th>Combined MBA Mean (Stddev) (n=110)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Radar graphs are most useful in assessing key profitability, liquidity, efficiency and solvency trends over several years.</td>
<td>3.17b (1.061)</td>
<td>3.42b (1.273)</td>
<td>3.30 (1.162)</td>
</tr>
<tr>
<td>2. Numerical ratios are most useful in assessing key profitability, liquidity, efficiency and solvency trends over several years.</td>
<td>3.71b (0.825)</td>
<td>2.73b (0.931)</td>
<td>3.28 (1.024)</td>
</tr>
<tr>
<td>3. Radar graphs together with numerical ratios are most useful in assessing key profitability, liquidity, efficiency and solvency trends over several years.</td>
<td>4.04b (0.885)</td>
<td>4.67b (0.734)</td>
<td>4.39 (0.858)</td>
</tr>
</tbody>
</table>

*a Five-point Likert scale was used for items one to three, the following endpoints were used: 1= strongly disagree, 5= strongly agree.

*bStatistical significance based on paired t-tests. Because of potential violations of assumptions for the parametric t-tests used, we also applied nonparametric tests, which yielded very comparable results. All mean differences have a p-values are <.001 except question #1, P < 0.10.
Table 5: Classroom Assessment Data
Student’s Self-Assessment of Knowledge/Impact of Radar Graphs (n=110)

<table>
<thead>
<tr>
<th>Response Item (For each of the knowledge impact statements, rate your level of agreement):</th>
<th>Regular MBA Mean (Stddev) (n=53)</th>
<th>Executive MBA Mean (Stddev) (n=57)</th>
<th>Combined MBA Mean (Stddev) (n=110)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Radar graphs have strongly affected my ability to learn ratio analysis because pictures make it easier than numbers.</td>
<td>3.63(^b) (0.908)</td>
<td>4.23(^b) (0.877)</td>
<td>3.94 (0.927)</td>
</tr>
<tr>
<td>2. During group presentations, I feel radar graphs made more of an impact on me than the ratios in numbers that were displayed and tabulated and presented.</td>
<td>4.00(^b) (0.970)</td>
<td>4.46(^b) (0.896)</td>
<td>4.23 (0.955)</td>
</tr>
<tr>
<td>3. I found it easier to segregate bad firms from good firms when using radar graphs in my group project firm analysis.</td>
<td>4.12(^b) (0.832)</td>
<td>4.62(^b) (0.889)</td>
<td>4.38 (0.878)</td>
</tr>
<tr>
<td>4. During group presentations, I feel that radar graphs enhanced the efficiency in which I spotted key year-over-year financial trends compared to the number ratios that were tabulated.</td>
<td>4.12(^b) (0.732)</td>
<td>4.75(^b) (0.622)</td>
<td>4.45 (0.737)</td>
</tr>
<tr>
<td>5. Radar graphs make it easier to interpret and summarize firm performance than using numerical ratios alone.</td>
<td>4.15(^b) (0.872)</td>
<td>4.73(^b) (0.660)</td>
<td>4.46 (0.809)</td>
</tr>
<tr>
<td>6. Radar graphs help keep me focused and interested when trying to interpret firm performance across multiple dimensions.</td>
<td>4.04(^b) (0.885)</td>
<td>4.67(^b) (0.617)</td>
<td>4.36 (0.819)</td>
</tr>
</tbody>
</table>

\(^a\) A five-point Likert scale was used for items one to six, the following endpoints were used: 1= strongly disagree, 5= strongly agree.

\(^b\) Statistical significance based on paired t-tests. Because of potential violations of assumptions for the parametric t-tests used, we also applied nonparametric tests, which yielded very comparable results. All mean differences have a p-values are <.001.
Table 6: Cross-Walk Student Self-Assessments and Independent Learning Outcome Assessments (n=110)

<table>
<thead>
<tr>
<th>L.O.</th>
<th>Self-Assessment Statements (Select)</th>
<th>Mean Response</th>
<th>Independent Assessment Stated Learning Objective</th>
<th>Mean Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Radar graphs are most useful in assessing key profitability, liquidity, and efficiency and solvency trends over several years. <em>(Table 4, Question #1)</em></td>
<td>3.30</td>
<td>Subjects should use radar graphs to demonstrate their understanding of firm performance across several dimensions: profitability, liquidity, efficiency and solvency trends.</td>
<td>3.80</td>
</tr>
<tr>
<td></td>
<td>Radar graphs together with numerical ratios are most useful in assessing key profitability, liquidity, efficiency and solvency trends over several years. <em>(Table 4, Question #3)</em></td>
<td>4.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>I found it easier to segregate bad firms from good firms when using radar graphs in my group project firm analysis. <em>(Table 5, Question #3)</em></td>
<td>4.38</td>
<td>Subjects should use radar graphs to demonstrate their ability to clearly segregate firm performance as either good or bad, as appropriate.</td>
<td>4.06</td>
</tr>
<tr>
<td>3</td>
<td>During group presentations, I feel that radar graphs enhanced the efficiency in which I spotted key year-over-year financial trends compared to the number ratios that were tabulated. <em>(Table 5, Question #4)</em></td>
<td>4.45</td>
<td>Subjects should use radar graphs clearly and concisely to assess firm performance on a single- and multi-year basis.</td>
<td>3.94</td>
</tr>
<tr>
<td>4</td>
<td>Radar graphs have strongly affected my ability to learn ratio analysis because pictures make it easier than numbers. <em>(Table 5, Question #1)</em></td>
<td>3.94</td>
<td>Subjects should use radar graphs to provide meaningful cross-sectional discussion and analysis regarding overall firm and industry trends.</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td>Radar graphs make it easier to interpret and summarize firm performance than using numerical ratios alone. <em>(Table 5, Question #5)</em></td>
<td>4.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Radar graphs make it easier to interpret and summarize overall firm performance than using numerical ratios alone. <em>(Table 5, Question #6)</em></td>
<td>4.36</td>
<td>Subjects should use radar graphs to demonstrate their understanding of each firms’ overall financial performance in a meaningful and persuasive fashion.</td>
<td>4.02</td>
</tr>
<tr>
<td></td>
<td>Radar graphs are more effective in synthesizing key performance trends compared to numerical (tabular) ratios alone. <em>(Table 3, Question #4)</em></td>
<td>4.24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*A five-point Likert scale was used for the Student Survey Scale, the following endpoints were used: 1= strongly disagree, 5= strongly agree.
*A five-point Likert scale was used for the Assessment Scale, the following endpoints were used: 1= did not meet expectations, 5= markedly exceeded expectations.

References
use. Accounting Horizons, 24, 257-278.
The Credit Quality Profile of The S&P 500 Index- A Historical Analysis

Dr. Vinod Venkiteshwaran
352 OCNR, College of Business
Texas A&M University
Corpus Christi TX 78411
vinod.venkit@tamucc.edu
Abstract
Despite the voluminous literature on the effects of additions and deletions to the S&P 500 index there exists no formal analysis of the overall credit quality or changes in the same for the S&P index constituents. Given that the index is a widely used measure of U.S. equities understanding the credit quality profile of the index constituents has important implications to both practitioners and academics. This study provides a first examination of the credit quality characteristics of the S&P 500 index constituents for the time period 1985-2006. The findings in the study suggests that over the 22 years considered the S&P index is increasingly composed of rated firms especially those with investment grade ratings of A and BBB. Further the index constituents that are not rated are typically smaller firms but with low leverage and high profitability compared to the index constituents that are rated.

Introduction
The Standard and Poor’s 500 (S&P) index is a widely followed market index of U.S. equities and is used as the basis for several index and exchange traded funds. Investment and asset management companies typically use the returns on the S&P 500 index to benchmark their own performance. S&P’s documentation on the selection criteria for the index stresses on the objective of reflecting the risk and return characteristics of the broader group of large capitalization equities. To achieve this end the criteria upon which index constituents are determined or changed include minimum market capitalization, float, profitability and liquidity. However these criteria do not stipulate any specific credit quality standards for the constituent firms. In other words index member firms are not constrained to be, for example, of investment grade quality or even have outstanding rated debt. Numerous researchers have examined the impact of additions or deletions to this index on market valuations, liquidity, the demand curve for stocks and other factors and yet no systematic evaluation of the credit quality of the index firms exists in the literature. This study examines the credit quality profile of the S&P 500 index firms over a 22 year time frame between 1985 and 2006. The empirical analysis is performed using aggregated credit ratings of the index member firms as well as standard determinants of credit quality as measured by firm specific financial characteristics.

Summary of results
- The analysis shows that a significant portion of the firms in the index are not rated but this trend is declining over the sample period. At the beginning of the sample period more than 50% of the index member firms are not rated but this proportion has reduced to less than 15% at the end of the sample period.
- Further the analysis indicates that, over the sample period, the index increasingly includes firms that are of investment grade quality. While less than 50% of the member firms carried an investment grade rating in 1986 more than 70% of the member firms are investment grade firms in 2006. However the proportion of firms with the top two investment grade ratings of AAA and AA have declined from 20% in 1986 to less than 9% in 2006. The proportion of the investment grade is dominated by a steady increase in

---

36 The S&P documentation for the listing criteria can be found at www.standardandpoors.com.
37 The listing criteria does not include any explicit credit quality standards but it does mention that firm leverage should be ‘operationally justifiable’.
38 See Denis, McConnell, Ovtchinnikov and Yu(2003), Chen and Singal (2004), Harris and Gurel (1986), Hegde and McDermott (2003), Elliott and Warr (2003), Lynch and Mendenhall (1997) and others
the proportion of firms with the bottom two rating category of investment grade ratings, A and BBB.

- Amongst the rated index constituents rating transition analysis shows that majority of the firms display very stable ratings. Fewer than 10% of the rated firm experiences a rating change while listed on the index. For the firms that do experience a rating change it is usually within one rating notch.

- Firm characteristics which are considered to be determinants of credit quality such as firm size, leverage levels, profitability and equity risk have changed significantly between the pre-1995 and post-1995 time periods for the constituent firms in the S&P 500 index.

- Further index constituents with and without credit ratings also differ substantially on the financial characteristics that affect credit quality. However the lack of credit rating is compensated for with low leverage and high profitability for these firms.

I. Motivation and Background

While there are many equity market indices the S&P 500 index is perhaps the most widely used and studied measure of stock market performance. The principal objective of any equity market index is to capture movements in the overall stock market and the S&P fact sheet on the S&P 500 index alludes to the same. “The goal of the Index Committee is to ensure that the S&P 500 remains a leading indicator of U.S. equities, reflecting the risk and return characteristics of the broader large cap universe on an on-going basis”. This suggests that changes in the index and its constituents are meant to reflect the overall characteristics of the broader equity market participants. There is a vast body of literature that examines the impact of additions and deletions to this index on the market valuations and price pressure effects surrounding the changes to the index. The primary motivation behind using the S&P 500 sample for empirical studies is due to the fact that there is no apriori reason to believe that index additions convey new information about the company and therefore acts as a natural experiment to test various asset pricing issues. Denis et al (2003) find that addition to index conveys new information about the company that was not previously incorporated in stock prices. Chen et al (2004) in a related study find that while there is a permanent increase in the valuation of newly added firms to the index there is no corresponding decline in the valuation of the firms that are removed from the index. Similar studies exist on the liquidity of the stocks post-addition to the index. See Hegde and McDermott (2003).

Despite the voluminous nature of the research on the index there exists no formal analysis of the credit quality of the index constituents. This aspect of the index composition is important for a variety of reasons. Firstly, if the S&P 500 index is indeed a representation of the broader equity markets then the credit quality profile of the index is a reflection of the credit quality of U.S. firms in general. Secondly, since the S&P 500 index is a widely used benchmark to evaluate investment performance understanding the risk profile of the underlying firms is important to portfolio managers and investors in general. And finally, since S&P does not explicitly account for credit quality in their listing criteria any systematic changes in the credit quality characteristics of the index constituents is an important consideration for academic studies that use the S&P index as an experimental data set as well.

II. Data and Results

The data used in this study comes from the COMPUSTAT database and includes all constituent firms of the S&P 500 index for the time period 1985-2006 with non-missing data. The sample
only includes those firms that are included in the S&P 500 index and thus any time a firm is delisted from the index it is automatically eliminated from the sample.

**Credit quality measured by credit ratings**
The sample is first analyzed for the proportion of index constituents that have a long term issuer credit rating as assigned by S&P. The long term issuer credit rating is S&P’s opinion of the overall credit worthiness of the firm. Prior to 1998 this variable reflects the rating assigned to a firm’s senior debt issue or to its subordinated debt when senior rated debt is not present. Therefore observations prior to 1998 with no issuer level credit rating imply that those firms did not have a rated debt issue. The sample is then classified on the basis of rating assignments as shown in Table 1. The S&P credit ratings go from high to low are AAA, AA, A, BBB and so on. All ratings better than a BBB are considered to be of investment grade quality and ratings below that are considered to be junk or speculative grade.

Panel A of Table 1 shows the frequency distribution for each year in the sample across the different rating categories and the distribution of those index constituents that did not have a long term issuer credit rating. Panel B shows the same frequency distribution as a percentage of the annual sample size. For convenience the same information is presented graphically in Figure 1 and Figure 2 as well. Figure 1 shows the overall sample composition for the entire sample period and Figure 2 shows the distribution on an annual basis. For the overall sample the majority of the firms have an investment grade rating of BBB or better but almost a third of the sample does not have a credit rating and a small proportion of firms have speculative grade ratings. The overarching pattern in Figure 2 is that the proportion of firms that do not have a long term credit rating has steadily declined over the sample period while the proportion of investment grade ratings has increased. This could be partly due to S&P assigning firm level ratings more systematically after 1998. However amongst the investment grade constituents the bottom two ratings of A and BBB seem to gain dominance toward the latter half of the sample period while there is a sharp decline in the proportion of firms with the AAA and AA ratings. If the index is indeed a reflection of the broader market this finding implies a decline in credit quality of the average large capitalization U.S. firm.

**Rating transitions of index constituents**
The results presented thus far indicate that the credit quality profile of the S&P index has indeed morphed over time with increasing proportions of rated firms and an increased concentration of investment grade firms. It is possible that the patterns of credit quality as measured by the credit ratings of the index constituents is due to rating transitions as opposed to the inclusion of new firms with credit ratings to the index. That is firms that are already in the index undergo rating upgrades and thus resulting in the observed trends as opposed to deliberate attempts by S&P to add highly rated firms to the index. To examine this possibility rating transitions, changes in credit ratings, of the index constituents is analyzed. This analysis is presented in Table 2.

Before discussing the results of the rating transition analysis it should be noted that 70% of the index member firms remain in the sample for the entire sample period therefore if inclusion of new firms is responsible for the credit quality profile changes of the index then this is due to a minority of the sample. Table 2 shows the changes in rating for each rating class included in the sample along with the frequency of rating changes measured as the number of notches over which the change occurred. For example looking at the row for the AAA rating 243 of the 247 firms with that rating experience no change in their long term credit rating. But 4 out of the 247 firms were upgraded from an AA rating (positive rating change of ‘1’ notch). The overall results in this table show that a majority, over 90%, of the index constituents experienced no rating
change and about 8% experienced a rating change of one notch. These findings suggest that rating transition by itself is not responsible for the general upward trend observed in the proportion of rated firms in the index and especially those with investment grade ratings.

Credit quality measured using firm financial characteristics

In this section credit quality is examined using firm specific financial characteristics that the extant literature has identified as determinants of credit quality. See, Blume, Lim and Mackinlay (1998) and Amato and Furfine (2005). The variables that are considered to be important factors of credit quality are firm size, profitability, leverage and equity risk. Firm size can be considered to be a measure of diversification or economies of scale and therefore is expected to be positively associated with credit quality. Measures of profitability such as interest coverage ratios and operating margins demonstrate the ability to service debt and hence are also considered to be positive determinants of credit quality. Leverage and equity risk, measured as the market model slope, on the other hand are expected to negatively affect credit quality. Therefore these characteristics are examined for index constituents across the sample period and also across firms with and without credit ratings. The results of this analysis are reported in Table 3.

Panel A presents the full sample means of the aforementioned financial characteristics and also the difference in mean t-statistics for these variables across the sub-sample periods before and after 1995. The t-statistics indicate that all the financial variables, except for total debt leverage, that capture credit quality are significantly different across these sub-samples. Index constituents are in general larger, have more long term debt, are more profitable and are less risky in the post 1995 time period.

Panel B examines the same financial characteristics across the two sub-sample periods between rated and unrated index constituents. Rated firms are fundamentally different from firms without ratings across all financial characteristics. The results suggest that unrated firms typically are smaller, have lower leverage and are more profitable compared to the rated firms. This pattern holds for both sub-sample periods before and after 1995. Therefore these results suggest that while the S&P index includes firms without credit ratings, less frequently in the latter half of the sample, the lack of a credit rating is made up for by better financial characteristics such as low leverage and higher profitability.

III. Conclusion

In this short study an examination of the credit quality profile of the S&P 500 index constituents over the 1985-2006 time period is presented. The analysis finds several interesting patterns in the data. The highlights of the findings are that the index is increasingly composed of firms with investment grade credit ratings and this trend is not driven by changes in the ratings of the index constituents but due to deliberate actions of S&P to include rated firms. Also the analysis of firm financial characteristics suggests that over time the S&P index includes larger firms with higher leverage and greater profitability. The results also indicate that unrated firms in the index while smaller have lower leverage and higher profitability compared to the rated index constituents. The analysis in this study has important implications to portfolio managers, investment management companies and academics alike.

References


Figure 1. Composition of S&P 500 constituents between 1985-2006

Figure 2. Composition of S&P 500 firms by year

- No Rating: 29%
- AAA: 3%
- AA: 10%
- A: 29%
- BBB: 20%
- BB: 7%
- B: 2%
- C: 0%
- CCC: 0%
Table 1. Composition of S&P 500 index constituents. This table presents the frequency distribution of the S&P 500 constituent firms on the basis of whether they carried a credit rating or not over the sample period 1985-2006. The table also shows the breakdown of the credit rating notch for those firms that did have a credit rating while listed on the index. The credit rating considered here is the long term issuer credit rating (COMPUSTAT data item 280/Xpressfeed data item splirc). Panel A presents the frequency distribution while Panel shows the same as percentages of annual sample sizes.

<table>
<thead>
<tr>
<th>Year</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>C</th>
<th>No Rating</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>11</td>
<td>55</td>
<td>61</td>
<td>24</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>179</td>
<td>341</td>
</tr>
<tr>
<td>1986</td>
<td>14</td>
<td>57</td>
<td>85</td>
<td>26</td>
<td>13</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>143</td>
<td>348</td>
</tr>
<tr>
<td>1987</td>
<td>16</td>
<td>57</td>
<td>90</td>
<td>29</td>
<td>17</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>135</td>
<td>355</td>
</tr>
<tr>
<td>1988</td>
<td>15</td>
<td>52</td>
<td>94</td>
<td>30</td>
<td>16</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>144</td>
<td>364</td>
</tr>
<tr>
<td>1989</td>
<td>15</td>
<td>53</td>
<td>92</td>
<td>39</td>
<td>17</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>143</td>
<td>369</td>
</tr>
<tr>
<td>1990</td>
<td>14</td>
<td>55</td>
<td>91</td>
<td>49</td>
<td>17</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>149</td>
<td>384</td>
</tr>
<tr>
<td>1991</td>
<td>13</td>
<td>55</td>
<td>98</td>
<td>46</td>
<td>16</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>164</td>
<td>401</td>
</tr>
<tr>
<td>1992</td>
<td>15</td>
<td>56</td>
<td>103</td>
<td>51</td>
<td>19</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>160</td>
<td>409</td>
</tr>
<tr>
<td>1993</td>
<td>13</td>
<td>54</td>
<td>118</td>
<td>51</td>
<td>24</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>158</td>
<td>420</td>
</tr>
<tr>
<td>1994</td>
<td>12</td>
<td>54</td>
<td>119</td>
<td>55</td>
<td>27</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>161</td>
<td>432</td>
</tr>
<tr>
<td>1995</td>
<td>12</td>
<td>49</td>
<td>130</td>
<td>69</td>
<td>29</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>150</td>
<td>444</td>
</tr>
<tr>
<td>1996</td>
<td>13</td>
<td>50</td>
<td>137</td>
<td>74</td>
<td>25</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>147</td>
<td>453</td>
</tr>
<tr>
<td>1997</td>
<td>12</td>
<td>47</td>
<td>145</td>
<td>90</td>
<td>29</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>137</td>
<td>464</td>
</tr>
<tr>
<td>1998</td>
<td>11</td>
<td>46</td>
<td>155</td>
<td>100</td>
<td>27</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>126</td>
<td>471</td>
</tr>
<tr>
<td>1999</td>
<td>12</td>
<td>42</td>
<td>161</td>
<td>114</td>
<td>35</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>111</td>
<td>479</td>
</tr>
<tr>
<td>2000</td>
<td>12</td>
<td>39</td>
<td>159</td>
<td>131</td>
<td>38</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>97</td>
<td>482</td>
</tr>
<tr>
<td>2001</td>
<td>11</td>
<td>38</td>
<td>153</td>
<td>150</td>
<td>41</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>85</td>
<td>487</td>
</tr>
<tr>
<td>2002</td>
<td>10</td>
<td>28</td>
<td>158</td>
<td>150</td>
<td>48</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>79</td>
<td>489</td>
</tr>
<tr>
<td>2003</td>
<td>10</td>
<td>27</td>
<td>157</td>
<td>150</td>
<td>52</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>74</td>
<td>489</td>
</tr>
<tr>
<td>2004</td>
<td>9</td>
<td>25</td>
<td>159</td>
<td>157</td>
<td>46</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>74</td>
<td>489</td>
</tr>
</tbody>
</table>
Panel B

<table>
<thead>
<tr>
<th>Year</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>C</th>
<th>No Rating</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>3.23%</td>
<td>16.13%</td>
<td>17.89%</td>
<td>7.04%</td>
<td>2.35%</td>
<td>0.88%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>52.49%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1986</td>
<td>4.02%</td>
<td>16.38%</td>
<td>24.43%</td>
<td>7.47%</td>
<td>3.74%</td>
<td>2.59%</td>
<td>0.29%</td>
<td>0.00%</td>
<td>41.09%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1987</td>
<td>4.51%</td>
<td>16.06%</td>
<td>25.35%</td>
<td>8.17%</td>
<td>4.79%</td>
<td>2.82%</td>
<td>0.28%</td>
<td>0.00%</td>
<td>38.03%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1988</td>
<td>4.12%</td>
<td>14.29%</td>
<td>25.82%</td>
<td>8.24%</td>
<td>4.40%</td>
<td>3.57%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>39.56%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1989</td>
<td>4.07%</td>
<td>14.36%</td>
<td>24.93%</td>
<td>10.57%</td>
<td>4.61%</td>
<td>2.71%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>38.75%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1990</td>
<td>3.65%</td>
<td>14.32%</td>
<td>23.70%</td>
<td>12.76%</td>
<td>4.43%</td>
<td>2.34%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>38.80%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1991</td>
<td>3.24%</td>
<td>13.72%</td>
<td>24.44%</td>
<td>11.47%</td>
<td>3.99%</td>
<td>2.24%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>40.90%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1992</td>
<td>3.67%</td>
<td>13.69%</td>
<td>25.18%</td>
<td>12.47%</td>
<td>4.65%</td>
<td>1.22%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>39.12%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1993</td>
<td>3.10%</td>
<td>12.86%</td>
<td>28.10%</td>
<td>12.14%</td>
<td>5.71%</td>
<td>0.48%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>37.62%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1994</td>
<td>2.78%</td>
<td>12.50%</td>
<td>27.55%</td>
<td>12.73%</td>
<td>6.25%</td>
<td>0.93%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>37.27%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1995</td>
<td>2.70%</td>
<td>11.04%</td>
<td>29.28%</td>
<td>15.54%</td>
<td>6.53%</td>
<td>1.13%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>33.78%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1996</td>
<td>2.87%</td>
<td>11.04%</td>
<td>30.24%</td>
<td>16.34%</td>
<td>5.52%</td>
<td>1.55%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>32.45%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1997</td>
<td>2.59%</td>
<td>10.13%</td>
<td>31.25%</td>
<td>19.40%</td>
<td>6.25%</td>
<td>0.86%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>29.53%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1998</td>
<td>2.34%</td>
<td>9.77%</td>
<td>32.91%</td>
<td>21.23%</td>
<td>5.73%</td>
<td>1.27%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>26.75%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1999</td>
<td>2.51%</td>
<td>8.77%</td>
<td>33.61%</td>
<td>23.80%</td>
<td>7.31%</td>
<td>0.63%</td>
<td>0.21%</td>
<td>0.00%</td>
<td>23.17%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2000</td>
<td>2.49%</td>
<td>8.09%</td>
<td>32.99%</td>
<td>27.18%</td>
<td>7.88%</td>
<td>1.04%</td>
<td>0.21%</td>
<td>0.00%</td>
<td>20.12%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2001</td>
<td>2.26%</td>
<td>7.80%</td>
<td>31.42%</td>
<td>30.80%</td>
<td>8.42%</td>
<td>1.44%</td>
<td>0.21%</td>
<td>0.21%</td>
<td>17.45%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2002</td>
<td>2.04%</td>
<td>5.73%</td>
<td>32.31%</td>
<td>30.67%</td>
<td>9.82%</td>
<td>3.27%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>16.16%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2003</td>
<td>2.04%</td>
<td>5.52%</td>
<td>32.11%</td>
<td>30.67%</td>
<td>10.63%</td>
<td>3.89%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>15.13%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Year</td>
<td>1.84%</td>
<td>5.11%</td>
<td>32.52%</td>
<td>32.11%</td>
<td>9.41%</td>
<td>3.89%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>15.13%</td>
<td>100.00%</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>2004</td>
<td>1.64%</td>
<td>4.91%</td>
<td>32.72%</td>
<td>33.74%</td>
<td>8.79%</td>
<td>3.27%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>14.93%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2005</td>
<td>1.64%</td>
<td>5.32%</td>
<td>31.49%</td>
<td>34.15%</td>
<td>10.84%</td>
<td>3.07%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>13.50%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Total</td>
<td>2.79%</td>
<td>10.36%</td>
<td>29.11%</td>
<td>20.08%</td>
<td>6.70%</td>
<td>2.05%</td>
<td>0.05%</td>
<td>0.01%</td>
<td>28.85%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>
Table 2. Rating transition analysis. This table presents the rating changes for the S&P 500 index constituents over the sample period 1985-2006. For each rating category, AAA-C ranked from 7 to 0, the frequency of rating changes is shown where the rating changes are represented as the number of notches over which the rating change occurred.

<table>
<thead>
<tr>
<th>Rating (Rank)</th>
<th>-5</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA (7)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>243</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>247</td>
</tr>
<tr>
<td>AA (6)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>866</td>
<td>34</td>
<td>0</td>
<td>912</td>
</tr>
<tr>
<td>A (5)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>82</td>
<td>2,447</td>
<td>77</td>
<td>0</td>
<td>2,609</td>
</tr>
<tr>
<td>BBB (4)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>128</td>
<td>1,594</td>
<td>67</td>
<td>1</td>
<td>0</td>
<td>1,794</td>
</tr>
<tr>
<td>BB (3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>61</td>
<td>477</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>573</td>
</tr>
<tr>
<td>B (2)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>22</td>
<td>131</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>165</td>
</tr>
<tr>
<td>CCC (1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>C (0)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>16</td>
<td>306</td>
<td>5,761</td>
<td>215</td>
<td>2</td>
<td>1</td>
<td>6,305</td>
</tr>
</tbody>
</table>

Table 3. Credit quality determinants- Differences across sample period and between rated and unrated index constituents. This table presents the means for various financial characteristics that are considered to be determinants of firm credit quality. The table also shows the difference in mean t-statistics and their significance levels across the sample period, which is split at 1995, and also between rated and unrated index constituents. In the following variable descriptions the COMPUSTAT data items and variable calculations are noted in parentheses. The Size variable is the log of book assets (at), LT debt is long term debt ratio(dltt/at), Total Debt is the total debt ratio (dlc+dltt/at), Coverage is the interest coverage ratio (oibdp/xint), operating margin is a profitability measure (oibdop/sale) and Beta is the market model slope parameter estimated over 200 trading days prior to the data observation date.

<table>
<thead>
<tr>
<th>Panel A. Pre versus Post-1995</th>
<th>Size</th>
<th>LT Debt</th>
<th>Total Debt</th>
<th>Coverage</th>
<th>Operating Margin</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Sample</td>
<td>8.48</td>
<td>0.19</td>
<td>0.24</td>
<td>23.90</td>
<td>0.22</td>
<td>1.01</td>
</tr>
<tr>
<td>Pre-1995</td>
<td>7.83</td>
<td>0.18</td>
<td>0.24</td>
<td>20.63</td>
<td>0.21</td>
<td>1.06</td>
</tr>
<tr>
<td>Post-1995</td>
<td>9.01</td>
<td>0.19</td>
<td>0.25</td>
<td>26.60</td>
<td>0.22</td>
<td>0.98</td>
</tr>
<tr>
<td>Difference in Mean t-stat(Post-Pre)</td>
<td>33.37</td>
<td>4.29</td>
<td>1.39</td>
<td>3.72</td>
<td>6.38</td>
<td>-6.70</td>
</tr>
<tr>
<td>P-value (2-tail)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1643</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Rated versus Not Rated</th>
<th>Size</th>
<th>LT Debt</th>
<th>Total Debt</th>
<th>Coverage</th>
<th>Operating Margin</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rated</td>
<td>9.06</td>
<td>0.21</td>
<td>0.27</td>
<td>12.21</td>
<td>0.21</td>
<td>0.96</td>
</tr>
<tr>
<td>Not Rated</td>
<td>7.01</td>
<td>0.13</td>
<td>0.17</td>
<td>58.91</td>
<td>0.22</td>
<td>1.18</td>
</tr>
<tr>
<td>Difference in Mean t-stat(Rated-Not Rated)</td>
<td>53.23</td>
<td>20.35</td>
<td>22.18</td>
<td>-16.20</td>
<td>-2.56</td>
<td>-13.84</td>
</tr>
<tr>
<td>P-value (2-tail)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0106</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>Rated</td>
<td>Not Rated</td>
<td>t-stat(Rated-Not Rated)</td>
<td>P-value (2-tail)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-------</td>
<td>-----------</td>
<td>-------------------------</td>
<td>------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.55</td>
<td>0.20</td>
<td>0.27</td>
<td>8.62</td>
<td>0.20</td>
<td>1.04</td>
</tr>
<tr>
<td>Difference in Mean</td>
<td>6.68</td>
<td>0.15</td>
<td>0.20</td>
<td>40.94</td>
<td>0.22</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Rated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference in Mean</td>
<td>34.99</td>
<td>8.86</td>
<td>10.51</td>
<td>-11.65</td>
<td>-5.41</td>
<td>-2.91</td>
</tr>
<tr>
<td>P-value (2-tail)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0037</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-1995</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Rated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference in Mean</td>
<td>36.39</td>
<td>21.28</td>
<td>23.73</td>
<td>-12.64</td>
<td>-0.14</td>
<td>-14.89</td>
</tr>
<tr>
<td>P-value (2-tail)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.892</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
The Impact of Mortgage Securitization on the Housing Bubble and Subprime Mortgage Crisis: A Self-organization Perspective

Hong Y. Park
College of Business and Management
Saginaw Valley State University
University Center, MI 48710, U.S.A.
hyp@svsu.edu

Hyejung Chang
School of Management
Kyung Hee University
Seoul, Korea
hjchang@khu.ac.kr

Kaustav Misra
College of Business and Management
Saginaw Valley State University
University Center, MI 48710, U.S.A
kmisra@svsu.edu

This research is funded by SVSU faculty research grant and the first author acknowledges Professor Diane Boehm for her editorial help for many years.
Abstract
The subprime mortgage crisis has been analyzed from many different perspectives. The securitization of subprime mortgages has emerged as the leading cause of the subprime mortgage crisis. This securitization is a complex process that involves a number of different players (Ashcraft and Schuerman, 2008). Securitization of subprime mortgages, which are a part of mortgage-backed securities (MBSs) also led to further complexity by the introduction of collateralized debt obligations (CDOs) and credit default swaps (CDSs). MBSs, CDOs and CDSs became sources of adverse selection and moral hazard which have contributed significantly to the current subprime mortgage crisis. Securitization of mortgages also made the mortgage market global, which provided opportunities for homebuyers in the U.S. to draw funds from all over the world. Our study investigates the impact of securitization of mortgages on mortgage rates, the housing bubble and the subprime mortgage crisis. The size of shadow banking grew significantly due to securitization of subprime mortgages in 1995.

We found that an application of the self-organization principle in biology to the analysis of the current housing bubble provides insight in the current subprime mortgage crisis and other bubbles in general. This study may offer a fresh new perspective for policy makers.

JEL classification: B52, G14, G21

Key words: subprime mortgage crisis, housing bubble, complexity, securitization of subprime mortgages, self-organization, opacity of mortgage financing, mispricing of risk, TARP, MBS, CDO, CDS, shadow banking

I. Introduction
The subprime mortgage crisis caused a major slowdown in the U.S. economy and economies around the world, imposing severe adverse effects on millions of people (Wu and Yang 2008). Numerous economists have analyzed the nature of the subprime mortgage crisis and offered solutions to the crisis. Shiller (2008) argues that the housing bubble, caused by irrational exuberance, led to the dangerous over-expansion of credit, which resulted in a global credit crunch. In a similar vein, Morris (2008) points out that the expansion of credit generated by home equity credit fueled high consumption and strong economic performances. During the housing bubble, credit was extended for “ninja” loans—no income, no job, no assets—and ninja loans became a part of subprime mortgage loans.

Some writers argue that the subprime mortgage crisis was caused by the housing bubble; others seek the origin of the crisis in the securitization of mortgages. It is likely that the housing bubble is related to the securitization of mortgages; thus treating the housing bubble and securitization of mortgages independently may not give a full picture of the current subprime mortgage crisis. Securitization of mortgages increases the liquidity of mortgages by creating secondary markets for them. As the improved liquidity of mortgages becomes a more desirable asset, demand for the asset increases. As demand increases, prices of mortgages increase and mortgage rates decline. The resulting housing prices and positive feedback of higher housing prices reinforce the housing bubble, demonstrating the relationship between the housing bubble and the securitization of mortgages.

The securitization of mortgages transformed the “originate and hold” mortgage model to the “originate and distribute” model (Mizen, 2008, p. 538). The originating bank holds the mortgage until the mortgage is fully paid off or sold to a new buyer and bears the full risk of mortgage defaults and rising mortgage rates in the originate and hold model. Therefore, the originating
bank has a strong incentive to screen the mortgage loan applicants carefully and monitor them after the mortgage deal is closed. The mortgages remain at the local banks. The securitization of mortgages distributes mortgages worldwide. The originate and distribute model has, however, contributed to mispricing of risk (Mizen, 2008), moral hazard and adverse selection (Ashcraft and Schuermann, 2008). The collapse of the housing bubble and mispricing of risk appear to be key culprits of the subprime mortgage crisis. This paper, therefore, examines the process of the transformation of the originate and hold model to the originate and distribute model, and the relationship between the housing bubble and the securitization of mortgage credit.

We can draw an analogy between the formation of the housing bubble and the self-organization principle in biology and thermodynamics. Self-organization in biological systems refers to a broad range of pattern-formation processes in both physical and biological systems (Camazine et al., 2003). The formation of the bubble pattern in the housing industry can be analyzed from this self-organization perspective, which provides a useful analytical framework to explain the housing bubble pattern and offer a solution to the subprime mortgage crisis. The paper examines the origin and evolution of mortgage securitization and the impact of securitization of mortgages on the housing bubble and subprime mortgage crisis. We also discuss solutions to the subprime mortgage crisis from a self-organization perspective and an alternative to self-organization.

II. Securitization of Mortgages: Its Complexity and Impact

II.1 Transformation to the originate-to-distribute model

Securitization of subprime mortgages is a result of the transformation of mortgage financing from the originate and hold model to the originate and distribute model. The complexity of this model grew over the years as financial institutions attempted to solve problems stemming from mortgage securitization.

This history may be traced to 1968, when the Government National Mortgage Association (Ginnie Mae) securitized Federal Housing Administration and Veterans Administration (FHA/VA) mortgages backed by the “full faith and credit” of the U.S. government for resale in the secondary market (Mizen, 2008, p. 536). Mizen (2008) indicates that government sponsored enterprises (GSE) such as the Federal National Mortgage Association (Fannie Mae) and Freddie Mac then began to securitize prime mortgages in 1980. Ginnie Mae is a government agent and Fannie Mae and Freddie Mac are government sponsored enterprises (GSEs). Securitization by a government agent and the GSEs meant that prime mortgage products and securitized prime mortgage products were subject to almost zero default risk. In contrast, private sector financial institutions’ involvement in securitization, including high quality (prime) loans, subprime loans and Alt-A loans and their MBSs, are subject to significant default risk (Mizen, 2008). In 1984 securitization of prime mortgage loans by private sector financial institutions emerged and in 1995 private financial institutions began to securitize subprime and Alt A mortgages (see Appendix A and B).

By 2006 according to Rosen (2007), Ginnie Mae’s guaranteed mortgages accounted for 4% of all mortgage backed securities (MBSs) issued. The GSEs involved 40% of MBSs; the remaining 56% were repackaged by private sector financial institutions. The mortgage backed securities (MBSs) share of total mortgage debt outstanding was about 56%, the non-MBSs share about 44%. MBSs, a kind of asset-backed securities (ABS), became more complex as private sector financial institutions issued more complicated new products, and pools of MBS were collected and securitized. Various special purpose vehicles (SPVs) were established to create new asset-backed securities from complex mixtures of residential MBSs, credit card and other debt
receivables; to avoid banks’ capital requirements, special purpose vehicles (SPVs) were treated as off-balance sheet items. Most MBSs included securities backed by prime loans, subprime loans or Alt-A loans, which are “issued to borrowers that appear to have good credit, but these loans do not meet the definition of prime or conforming” (Rosen, 2007, p. 2).

In this process, private financial institutions categorized asset-backed securities (ABS) into three tranches: senior, mezzanine, and equity levels based on the priority of the claim that holders of these financial instruments can make in case of bankruptcy. Bonds that are themselves backed by pools of bonds are referred to as collateralized debt obligations (Rosen, 2007); a number of the collateralized debt obligations (CDOs) purchased MBSs and securities of other CDOs. Banks hold asset-backed securities in warehouses before intermediating credit to end investors. Securitization gets complex as financial institutions develop a market for collateralized debt obligations, and Mizen’s (2008) statement on the complexity of securitization illustrates the process of complexity in mortgage securitization:

Some tranches of CDOs were then pooled and resold as CDOs of CDOs (the so-called CDOs-squared); CDOs-squared were even repackaged into CDOs-cubed (Mizen, 2008, p. 538).

These CDOs were distributed to final investors and various entities including primary lenders, mortgage brokers, bond insurers and credit rating agencies (OECD, 2008), all of whom participated in this process at various stages from origination to final distribution. Unlike the originate-to-hold model, which does not involve these stages and thus generates less credit risk, the originate-to-distribute model involves credit risk at each stage, and this mispricing of credit risk has been a key element in the current subprime mortgage crisis.

II.2 Problems in the originate-to-distribute model

What are the sources of credit risk? Economists contribute the mispricing of risk to information asymmetry and frictions stemming from it. According to Ashcraft and Scheurmann (2008), the securitization process is subject to seven key frictions:

1. Frictions between the mortgagor and the originator: predatory lending because subprime borrowers can be financially unsophisticated.
2. Friction between the originator and the arranger: predatory borrowing and lending; the originator has an information advantage over the arranger with regard to the quality of the borrower.
3. Frictions between the arranger and third-parties: adverse selection; the arranger has more information about the quality of the mortgage loans, which creates an adverse selection problem: the arranger can securitize bad loans (the lemon) and keep the good ones.
4. Frictions between the servicer and the mortgagor: moral hazard.
5. Frictions between the servicer and third-parties: moral hazard.
7. Frictions between the investor and credit rating agencies: model error (Ashcraft and Scheurman, 2008, pp. i and ii).

Because of these frictions, the originate-to-distribute model of mortgage lending creates opportunities for multiple problems: predatory lending and borrowing, moral hazard, adverse selection, principal-agent problem, and model error in credit rating. These problems led to mispricing of risk, which together with the housing bubble, caused the subprime mortgage crisis (Ashcraft and Scheurmann, 2008; Mizen, 2008; Morris, 2008; Rosen, 2007; Schiller, 2008). Since a significant portion of banks’ revenues are generated by fees on originating mortgages, servicing mortgages and issuing MBSs, these fee-generating activities have changed the nature
of banking. Banks began to engage in more fee-generating activities as they serviced mortgages and issued MBS. Therefore, securitization of mortgages likely fostered changes in banking practice to fee-generating banking. Banks continue to be involved in various fee generating activities such as ATMs and automatic loans.

Furthermore, the originate-to-distribute model creates a larger number of steps and opacity of the financial system between the originator and the final holder of mortgages. Mizen cites comments from Alexander Lamfalussy and William Buiter, the former general manager of the Bank for International Settlements and former chief economist of the European Bank for Reconstruction and Development, respectively, who note that “banks have replaced the ‘originate and hold’ model of lending long and borrowing short, with an ‘originate and distribute’ model in which they lend and then sell the claims to someone else” (Mizen, 2008, 550). The originate-to-distribute model was designed to solve problems stemming from lending long and borrowing short in the originate-to-hold model. However, this model has accompanied problems which were not solved during its development. Mizen (2008) argues that a larger number of steps between the originator and holder added greater opacity to the process and contributed to the mispricing of risk that was not properly appraised. He further points out that the extension of originate and distribute banking to subprime mortgage securities created an asset class with an opaque ownership structure, and is ultimately responsible for the subprime mortgage crisis, as banks created an asset class of special investment vehicles (SIVs) and put them on the off-balance sheet to avoid their capital requirements. According to Mishkin (2013), avoiding regulation is a typical behavior of financial institutions, but the consequences are often overlooked.

The asset class of special investment vehicles (SIVs) is placed on the off-balance sheet to reduce a risk-weighted-capital adequacy requirement. This is a shadow banking activity. Shadow banks are defined as financial intermediaries that conduct functions of banking without access to central bank liquidity or public sector credit guarantees (Adrian et al, 2010; Noeth and Sengupta, 2011). The shadow banking system consists of hedge funds, investment banks, and other nondepository financial firms Mishkin, 2013). The size of the shadow banking sector was close to $20 trillion in 2007 compared to $11 trillion of the traditional banking system. The size of shadow banking system began to increase sharply in 1995 which was the year that the securitization of subprime mortgages was initiated. The traditional banking system is insured by the Federal Deposit Insurance Corporation (FDIC) for banks’ runs. However, the shadow banking activities are not insured by the public sector and create crisis in the case of runs of the shadow banking system.

Declines in housing values accelerated runs of the shadow banking system because of the collapse of the housing bubble. Economists argue that the runs on the shadow banking system contributed to the current subprime mortgage crisis because investors withdraw their funds even at the smallest hint of bad news (Adrian and Shin, 2009; Shin, 2009). Many hedge funds participants withdrew funds at the onset of the current financial crisis. Hedge funds are subject to withdraw at the discretion of hedge fund participants and there was a sharp decline in hedge funds at the onset of the crisis. Loans of shadow banking systems are financed by repurchase agreements (repos) and concerns on the quality of the balance sheets in shadow banking systems led lenders require larger amounts of collateral (securities) which is known as haircuts (Adrian and Shin, 2010; Mishkin, 2013). As a result the availability of credits declined significantly, a liquidity shock was created and the economic decline followed.

According to Adrian, Kimbrough and Marchioni (2011) considerable strains in the commercial
paper market emerged following the bankruptcy of Lehman Brothers Holdings Inc. on September 15, 2008. As a result of the bankruptcy, money market investors reallocated their funds from prime money market funds to those that held only government securities. They state that as demand for money market funds shrank, commercial paper issuers were unable to issue term paper and instead issued overnight paper. Institutions relying on commercial papers became vulnerable to bankruptcy as money market fund investors pulled away from the commercial paper market. The liquidity of financial institutions dried up and business and commercial loans declined. As a result the U.S. economy fell into a deep recession.

II.3 Problems in managing mortgage credit risk
A further complexity is added to the originate-to-distribute model (Mizen, 2008) in management of mortgage credit risk: a credit default swap (CDS) is utilized for management of credit risk. The CDS is a credit derivative contract between two counterparties: the buyer makes periodic payments (premiums) to the seller, and in return receives a payoff (protection) if any underlying financial instrument defaults during the term of the CDS contract. CDSs can be bought by most institutional investors, but it is not necessary for the buyer to own any CDO. The cost of insurance to cover default risk using CDSs had become much more expensive as subprime mortgage default increased and the ABX declined during 2008.

The ABX index launched in January 2007, and is used as an indicator of default risk. The ABX index serves as a benchmark of the market for securities backed by home loans issued to borrowers with weak credit. The ABX indices are based credit derivatives written on MBS backed by subprime mortgage loans and track the price of credit default insurance on a basket of such deals (Fender and Scheicher, 2008). The ABX index tracks the performance of a basket of credit default swaps (CDSs) based on U.S. subprime loans, and traders and investors are allowed to take positions without actually holding CDSs (Wong, 2008). Sellers of CDS were not anticipating the collapse of the housing bubble and there were also misleading or fraudulent opportunistic positions taken by financial institutions. Financial institutions on the wrong sides of positions and sellers of CDS became insolvent as the housing bubble popped. These financial institutions were bailed out by the government. According to the Bank for International settlements report, there was an estimated $62.2 trillion worth of CDS contracts outstanding worldwide in 2008 (Morgensen, 2008-02-17, New York Times, ISDA market survey), but fell to $26.3 trillion by mid-year 2010. The originate-to-distribute model was designed to solve the problems in the originate-to-hold model of high interest rate risk and low liquidity in lending long and borrowing short (maturity transformation). However, the illiquidity and high interest rate risk of the originate-to-hold model led to high mortgage rates (qualitative asset transformation), which reduced housing demands. The problems of interest rate risk of lending long and borrowing short (the problem of qualitative asset transformation) became more severe in the early 1980s because of wide swings in short-term interest rates. However, the new solution came with new problems, such as increased opacity and mispricing of risk in the financial system. Problems generated by this new model have still not been adequately addressed.

Popper’s (1982) fundamental evolutionary sequence of events illustrates that solutions eliminate errors, but also generate new problems that need to be solved. Therefore, “all organisms are constantly, day and night, engaged in problem-solving” (Popper, p. 110). Policy makers might be advised to adopt Popper’s constant problem-solving framework throughout their entire evolution of any new lending model or policy.

III. Securitization of Mortgages, Mortgage Rates and Housing Bubble
III.1 Securitization of mortgages and mortgage rates
Securitization in corporate finance is now pervasive (Gorton and Souleles, 2005). How do financial corporations carry out securitization? Corporations use special purpose vehicles (SPVs) in securitization and remove debt from the balance sheet and obtain a lower cost of capital. Gorton and Souleles (2005) argue that SPVs exist in large part to reduce bankruptcy costs. Financial institutions make strategic choices for their financing decisions: on-balance sheet or off-balance sheet. Off-balance sheet activities separate control rights for the business decisions from the financing decisions. The SPV is a legal entity created by a firm. According to Gorton and Souleles (2005) SPVs are essentially robot firms that have no employees, make no substantive economic decisions, have no physical location, and cannot go bankrupt. The sponsoring or originating bank creates the SPV for securitization of mortgages and helps improve the originating bank’s returns on assets by reducing the size of assets from its balance sheet and increasing returns from SPVs. Prices of an asset are determined by the demand and supply of the asset. Determining factors of the demand for mortgages are relative expected returns, taxes, liquidity and wealth (Mishkin, 2013). Securitization of mortgage credit increases the liquidity of mortgages and shifts the demand for mortgage to the right. As the demand for mortgages increases, the price will go up and mortgage rates will decline, due to an inverse relationship between prices of bonds and interest rates (Mishkin, 2013). The increasing price of mortgages lowers the mortgage rate. Based on the effect of liquidity on mortgage rates, we can formulate the following hypothesis: Hypothesis 1: Securitization of mortgage credit lowers mortgage rates.

III.2 Securitization of Subprime Mortgages and the Housing Bubble

The housing bubble can be further explained by the spontaneous responses of actors in the mortgage market. What was the triggering or tuning mechanism of the housing bubble? Based on the model used above, securitization of subprime mortgages was the likely source. The questions are: What comes after the triggering or tipping point in a bubble and how does a bubble form? Shiller’s irrational exuberance (2008) and Keynes’ animal spirit (1936) attempt to answer the question. We argue that the self-organization theory in biology and thermodynamics (Kauffman, 1993; Prigogine and Stenger, 1984) help explain the formation of the bubble. Shiller’s home price index shows that the current nominal housing bubble started in 1995 and ended in 2006 (Shiller, 2008). He argues that the housing bubble was caused by irrational exuberance and the social contagion of boom thinking. He stresses the feedback effect in producing speculative bubbles:

Psychological, epidemiological, and economic theory all point to an environment in which feedback of enthusiasm for speculative assets, or feedback of price increases into further price increases, can be expected to produce speculative bubbles from time to time. They make clear that these bubbles can have complicated—sometimes random and unpredictable—dynamics. (Shiller, 2008, p. 47)

Akerlof and Shiller (2009) argue a similar point in their newly published book, Animal Spirits. They point out that a key to address the current problem is to recover Keynes’s (1936) insight about “animal spirits”—the spontaneous attitudes and ideas that guide economic action:

Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many years to come, can only be taken as a result of animal spirits of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities. (Keynes, 1936, p. 161)
The guiding principles of animal spirits are relatively simple and spontaneous. The self-organization theory is similar to animal spirits in terms of responding to information only in the area immediately around you. We argue that the self-organization perspective (Kauffman, 1993; Prigogine and Stenger, 1984) can explain the formation and expansion of the housing bubble well. Self-organization refers to a broad range of pattern formation processes in both physical and biological systems (Camazine et al., 2003). Camazine et al. define self-organization:

Self-organization is a process in which a pattern at the global level of a system emerges safely from numerous interactions among the lower-level components of the system. Moreover, the rules specifying interactions among the system’s components are executed with only local information, without reference to the global pattern. (Camazine et al., 2003, p. 8)

Furthermore, “The multiplicity of interactions that characterizes self-organization systems emphasizes that such systems are dynamic and require continual interactions among lower-level components to produce and maintain structure” (Camazine et al., 2003 p. 8). Home buyers, banks and regulators react to information on increasing housing prices; home buyers buy houses anticipating the appreciation of housing prices; and banks make mortgage loans with the expectation that mortgage loans are secured because home values will continue to exceed mortgage values. These spontaneous interactions and cascades of events form a housing bubble, which expands until the bubble reaches its maximum.

The concept of feedback in biological and physical systems (Kauffman, 1993; Prigogine and Stenger, 1984) can add more insight to analysis of the housing bubble pattern. Camazine et al. (2003) explain the relationship between feedback and change.

Feedback can have two basic values: positive or negative. Feedback is positive if the recurrent influence reinforces or amplifies the initial change. The snowballing effect of positive feedback takes an initial change in a system and reinforces that change in the same direction as the initial deviation. Self-enhancement, amplification, facilitation, and autocatalysis are all terms used in positive feedback. (Camazine et al., 2003, p.17)

The amplifying nature of positive feedback means that it has the potential to produce destructive explosions or implosions in any process where it plays a role. How can such snowballing be kept under control? This is where negative feedback plays a critical role, providing inhibition to offset the amplification and helping to shape it into a particular pattern. (Camazine et al., 2003, p. 19)

Camazine et al. (2003) illustrate self-organization phenomena using a school of fish and a herd of reindeer. A school of fish maneuvers gracefully, with all its members moving in parallel in the same direction. The reason that fish do not run into each other is due to a negative feedback, which helps maintain equilibrium for a school of fish. Similarly, a housing bubble moves upward, with all home buyers moving in the same direction. In other words, the housing bubble can be explained by persistent housing price increases due to an extended duration of positive feedback. Shin (2008) and Adrian and Shin (2010) point out the feedback that financial intermediaries adjust their balance sheets in such a way that leverage is high during booms and low during bust so that leverage is procyclical.
This raises a question: What triggered the initial change in the housing bubble? We believe that the initiation of private financial institutions’ involvement in subprime mortgage credit securitization was a tipping point in the current housing bubble. As stated before, the subprime MBSs by private financial institutions grew rapidly after their inception in 1995. The increasing MBSs by private financial institutions made credit available to subprime mortgagors and created high demands for housing, thus increasing housing prices. Home buyers then responded to the appreciation of housing prices, and financial institutions responded to opportunities for generating fees in mortgage origination, servicing, issuing MBSs and CDOs, and underwriting mortgages and MBSs. Both home buyers and financial institutions reacted based on their local information, and the positive feedback of the initial change formed and amplified the housing price bubble. Thus we can observe that the housing bubble pattern emerged from numerous interactions among the lower level of components (e.g., home buyers), as seen in the self-organization of biology and physics (Camazine et al., 2003; Haken, 1977; Kauffman, 1993; Prigogine and Stengers, 1984). Interactions among home buyers and financial institutions are made based on local information and they are components of housing and financial markets. The pattern of the current housing bubble has emerged from interactions among home buyers and financial institutions.

What then made the housing bubble collapse and reverse itself? There are two factors: the housing bubble exhausted credit availability and household debts reached capacity limits (Morris, 2008). To maintain equilibrium negative feedback is required (Kauffman, 1993). Positive feedback accompanies negative feedback and all bubbles, therefore, eventually collapse (Schiller, 2008). Negative feedback offers an opportunity for error elimination (Plotkin, 1982; Popper, 1982). Errors in the securitization and housing bubble were mispricing of risk and lack of due diligence in all actors in the originate-to-distribute model, including government regulatory agents. As housing prices decreased, speculative downward demand reinforced a decline in housing prices, creating a gap between the value of the mortgage and the value of a house.

Based on the impact of securitization and self-organization on the housing bubble, we can draw the following hypothesis:

Hypothesis 2: The current housing bubble is triggered by securitization of subprime mortgages.

Hypothesis 3: Positive feedback in securitization of subprime mortgages amplifies the housing bubble and negative feedback reverses the housing bubble: positive and negative feedback make the pattern of the housing bubble sigmoidal.

IV. Model and Empirical Results
IV.1 Model
We can test these hypotheses with models. In the mortgage rate model (equation 1), we include the mortgage securitization ratio to test hypothesis 1; the inflation rate variable is included based on the Fisher effect on nominal interest rates. The GDP growth rate variable is used to capture the relationship between the demand for an asset and income. We expect that there is an inverse relationship between the mortgage rate and securitization ratio, and the inflation rate and GDP growth rate variables are expected to show a positive relationship with the mortgage rate variable.

The housing price model (equation 2) is specified to test hypothesis 2 and 3, as well as four theories that have been frequently cited as explanations for the current housing bubble. The
housing price model includes the mortgage yield rate, mortgage securitization, real GDP growth rate, building cost index, and a dummy variable for subprime mortgage securitization. Except for the building cost index variable, these variables are determining factors of the demand for housing.

Models include the following elements to reflect theories of the mortgage rate and housing price:

1. Securitization of subprime and Alt A mortgage loans: This securitization started in 1995, the current housing bubble started at the same time. Subprime mortgage securitization increased housing demand by offering mortgages to less qualified home buyers at lower mortgage rates.

2. The irrational exuberance hypothesis (Robert Shiller, 2008)

3. The low mortgage rate (interest rate) hypothesis: The Fed Chairman, Greenspan, maintained lower interest rates, which led to lower mortgage rates during the housing bubble. There is an inverse relationship between housing prices and mortgage rates.

4. Securitization of mortgages: Securitization of mortgages provides more funds for home buyers; mortgage-backed securities (MBSs) were sold all over the world and drew savings from the U.S. as well as other countries. Securitization may also lower mortgage rates.

We specified our models to reflect these theories and models as follows:

\[ \text{MRT} = f(\text{INF}, \text{MSR}, \text{RGDPG}) \]  
\[ \text{HPI} = f(\text{MSR}, \text{RGDPG}, \text{BCI}, \text{SUBP}) \]  

Where,  
- MRT: the 10 year mortgage yield rate  
- INF: the inflation rate  
- MSR: the mortgage securitization ratio (the ratio of securitized mortgages to total mortgage outstanding)  
- RGDPG: the real GDP growth rate  
- HPI: the housing price index  
- BCI: the building construction cost index  
- SUBP: the dummy variable for securitization of subprime mortgages (SUBP=0 before 1995 and SUBP=1 after 1995)

The mortgage rate model includes inflation rate to test the Fisher hypothesis and the real GDP growth rate to test the procyclical pattern of interest rates in addition to the percentage of mortgage securitization. The housing price model is specified to investigate the impact of mortgage securitization and initiation of subprime on the housing bubble. This model includes the real GDP growth rate and building cost index variables.

**IV.2 Data and empirical results**

Data used in estimation of models are from Shiller (2008), the 2008 Presidential Report and the 2008 mortgage market statistical annual (Inside Mortgage Finance Publication). Detailed sources of each variable are as follows:

- MRT, INF, RGDPG: the 2008 Presidential report  
- MSR: estimated by the author based on the 2008 mortgage market statistical annual (Inside Mortgage Finance Publication)
Table 1: Estimation Results

\[
\begin{align*}
\text{MRT} &= 16.20 - 0.157 \text{MSR} - 0.197 \text{RGDPG} - 0.0547 \text{INF} \\
\text{HPI} &= -503.29 + 275.40 \text{MSR} - 3.7625 \text{RGDPG} + 5.832 \text{BCI} + 28.389 \text{SUB}
\end{align*}
\]

<table>
<thead>
<tr>
<th>Test</th>
<th>Break Point</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chow</td>
<td>15</td>
<td>5</td>
<td>15</td>
<td>14.16</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

The numbers in parentheses are p-values.

These data are time series data from 1980 to 2007. We employed both a multiple regression estimation method and the SAS statistical package to estimate regression coefficients and Chow test statistics. Regression coefficients and p-values are presented in Table 1. The mortgage rate variable is deleted from the housing price model because of a strong correlation between the mortgage rate and the securitization of mortgage. The mortgage securitization variable reveals a statistically significant inverse association with the mortgage rate and a statistically significant positive association with the housing price variable. These results support the main thrust of our investigation on the subprime mortgage crisis. Detailed discussions are presented in the following section.

V. Empirical Analysis and Discussion

V.1 Securitization and the Mortgage Rate

We specified the mortgage rate model to find the impact of mortgage securitization on the mortgage yield rate. Independent variables included in the model are mortgage securitization ratio, the real GDP growth rate, and the inflation rate. Securitization reveals a statistically significant inverse association with the mortgage rate, which conforms to the expectation and supports hypothesis 1 that mortgage securitization, has an inverse association with the mortgage rate. This result also conforms the finding of Kolari et al (1998) that securitization of mortgage had moderate to strong effects on lowering mortgage yield spread. However, the real GDP variable has a statistically significant negative association with the mortgage rate, and the inflation variable does not show a statistically significant association with the mortgage rate variable. Signs of these two variables do not conform to the expectations of the theoretical model.

V.2 Subprime Mortgage Securitization, the Housing Bubble and Self-organization

Model (2) tests whether the subprime mortgage securitization was a triggering factor in the U.S. housing bubble; regression coefficients are estimated based on data from 1980 to 2008. We included additional variables such as the building construction cost index and real GDP growth rate variables in this model. The real GDP growth rate variable is not statistically significant and does not reveal the expected positive relationship with the housing price. The building cost index
The mortgage securitization variable shows an expected sign and is statistically significant. The dummy variable representing securitization of subprime mortgages reveals a positive sign, although it is not statistically significant at the conventional $\alpha$ level. To find out whether subprime mortgage securitization led to a structural change in the housing price, we conducted a Chow test. The Chow statistics show that the break point is the 15th observation (1995), which is the year that subprime mortgage securitization started. The Chow test statistics are statistically significant and the test results suggest that subprime mortgage securitization was a trigger point for the current U.S. housing bubble. Figure 1 confirms the breaking point for the structural change.

When we examined data from 1980 to 2007, we found that the current housing bubble started in 1995. The mortgage securitization ratio started to rise significantly in 1990 due to a sharp increase in both government sponsored enterprises (GSEs) and non-agency mortgage securities. Securitization of non-agency mortgage securities (private sector financial institutions) had started in 1984 with a securitization ratio of 19.91%; by 2007 the securitization ratio rose to 59.59%. The non-government agency mortgage securities were only 4.2% in 1984, but rose to 31.98% in 2007, which can be calculated from columns (2) and (3) in Appendix B. Securitization of mortgages increases the liquidity of mortgages and reduces the interest rate risk of lending long and borrowing short. The problem of lending long and borrowing short became more apparent in the early ‘80s because of the wilder swing of the short-term interest rates. The securitization rate of mortgages, a parameter which changes between 0 and 1, is comparable to a tunable parameter in biology (Camazine et al., 2003). There is no securitization of mortgage when the parameter is 0 and 100 percent securitization when it is 1. Changes in the securitization parameters result in changes in housing prices. Securitization is a strategic or policy variable in management or economics. Securitization allowed a bank to get the risky loans off its books and supposed to shift the risk to deep pocketed insurance companies and pension funds (Adrian and Shin, 2010; Rajan, 2010), but they ended up to banks because banks acquired each other’s risky assets to boost their returns. Therefore, securitization and its complexity contributed to the crisis.

What did trigger the current housing bubble? One answer may be securitization of subprime mortgage credits. Securitization of mortgage credit by Ginnie Mae started in 1968; other GSEs became involved in securitization in 1980. Securitization of subprime mortgages by private financial institutions started in the mid-1990s and increased rapidly. We argue that this is a tipping point. In thermodynamics and biology, the tipping point is referred to as the threshold point of bifurcation (Heylighen, 2008; Kauffman, 1993; Prigogine and Stengers, 1984). When a person raises the temperature, water does not boil until the temperature reaches the threshold point. Although securitization began earlier, the housing bubble did not start until the sharp increase in securitization of subprime mortgages in 1995. The snowballing effect of positive feedback takes an initial change in a system and reinforces that change in the same direction as the initial deviation (Camazine, et al., 2003, p. 17). In the same way, the initial housing bubble triggered by initial external subprime mortgage securitization in 1995 amplified the housing bubble and the pattern of housing prices changed in 1995.

To test the impact of subprime mortgage securitization on the housing bubble, we used a dummy variable for the year that the securitization of subprime mortgages was introduced. Empirical results show that this variable had a positive impact on the housing bubble. The downward snowballing effect of negative feedback can provide inhibition to offset the amplification. The onset of the decline in housing prices triggered by the subprime crisis (unsustainable debt) can be
amplified as time passes; home buyers act on information about increasing housing prices, they interact with each other, and they also interact with financial institutions. Consequently, the pattern of the housing bubble becomes nonlinear (see Figure 1).

The pattern of this housing bubble forms a sigmoidal curve. Kauffman’s (1993) explanation of a sigmoidal curve in self-organization can be applicable to the sigmoidal curve of the housing bubble pattern:

The sigmoidal function is initially below the proportional response. Here a given output levels to an output that is less than the input. Were that reduced output fed back as the next input, then the subsequent response would be even less. Over iterations, the response would dwindle to zero. The sigmoidal response becomes steep in its midrange, however, and crosses above the proportional response. An input above this critical crossing point leads to an output that is greater than the proportional-response output. In turn, were that output fed back as a next input, the output would be still greater than input. Over iterations the response would climb to a maximum. (Kauffman, 1993, p. 184)

As discussed earlier, the pattern of the housing bubble was formed by numerous interactions of home buyers, financial institutions and credit rating agencies. As long as a housing bubble continues to expand, MBSs are secured debts because they are asset-backed securities (ABSs). Securitization of subprime mortgages offers more mortgage loans to home buyers, and housing prices increase as a result of increased demand. As housing prices climbed to a maximum in 2006, the housing bubble reversed itself due to constraint of household debt capacity. These behaviors of housing prices make the pattern of the housing bubble sigmoidal, as we can see in Figure 1. Camazine et al. (2003) have shown how a small change in a system parameter can result in a large change in the overall behavior of the system (p. 35). Figure 1 shows that a change in securitization caused by subprime mortgage securitization resulted in large changes in housing prices. This also confirms the butterfly effect, the analogy that a butterfly flapping its wings could cause hurricanes in another part of the world (Lorenz, 1963; Rosser, 2000). Miller and Page (2007) also illustrate how systems of interacting agents can lead to emergent phenomena. Their linking individually based micro processes to macrosocial outcomes is a useful analogy to the formation of housing bubble phenomena. This result shows a nonlinear relationship between inputs and outputs. Thermodynamics and self-organization offer possibilities those small inputs can generate large outcomes and large inputs may result in small outputs; they provide a different perspective from classical Newtonian physics. In the Newtonian universe, outputs are proportional to inputs and the perturbation eventually gravitates toward equilibrium. However, in thermodynamics and self-organization, the perturbation reaches to a new attractor (equilibrium) by going through rugged landscapes. The originate-to-distribute model perturbed the originate-to-hold model and went through rugged landscapes. The new attractor (equilibrium) will have characteristics that are different from the originate-to-hold model. Policy makers and managers need to become aware of characteristics of the new attractor as they address problems stemming from the new order.

Securitization of home mortgages has continued to increase by adding more institutions, such as the government agency (Ginnie Mae), government sponsored enterprises (Fannie Mae and Freddie Mac) and private sector financial institutions. Securitization of private financial institutions includes prime mortgages, subprime mortgages and Alt A mortgages. The pattern of the housing bubble changes as securitization of mortgages increases. Figure 1 also illustrates the
nature of the relationship between self-organization and housing prices. This relationship is supporting evidence for hypothesis 3.

Figure 1: A Scatter plot of nominal housing prices and mortgage securitization

The maximum point is another bifurcation point where negative feedback reinforces the downward change or amplifies the initial change. In the case of a school of fish, positive feedback and negative feedback work with little time lag, so that each fish maintains its distance with neighboring fish and equilibrium (Camazine et al., 2003). However, the housing bubble pattern shows that positive feedback amplifies a housing bubble and controlling or negative feedback operates with a significant time lag. The housing bubble started in 1995 and lasted until 2006. Although operating rules for home buyers and financial institutions are simple, the global pattern of spontaneous interactions of home buyers and financial institutions became complex. This complex global pattern of spontaneous interactions led to the crisis. Data for Figure 1 are from Shiller (2008) and the authors’ estimation based on the Inside Mortgage Finance Report (2008).

V.3 Securitization of Subprime Mortgages and Subprime Mortgage Crisis

Column (5) in Appendix B shows the share of the subprime and Alt A mortgages out of total non-agency issuance (private financial institutions). The share of the subprime and Alt A mortgages has been growing since 1995 and peaked in 2006, with 78.8% of the total MBS issuance amount by private financial institutions.

When securitization of subprime and Alt A loans by private sector financial institutions started in 1995, mortgage backed securities became a significant part of the collateralized debt obligations (CDOs). Resecuritization of CDOs and a derivative of CDOs, credit default swaps (CDSs), further complicated the subprime mortgage problem. As stated before, there was an estimated $62.2 trillion worth of CDSs contracts outstanding worldwide in 2006 (Morgensen, 2008).

Private sector financial institutions’ involvement in subprime mortgage securitization caused subprime loans and Alt A loans to increase due to the transferability of risk in the originate and
distribute model. Securitization of subprime complicates the mortgage financing and lengthens mortgage intermediation chains. Shin (2009) points out that long intermediation chains have been associated with the rapid development of the securitized, market-based financial system in the United States. Shin (2009) argues that “long intermediation chains carry costs in terms of greater amplitude of fluctuations in the boom and bust cycle of leverage and balance sheet size. Shorter intermediation chains carry benefits for stability of the financial system.” (p. 28). Thus, securitization of subprime loans and Alt A loans became a main source of the current subprime mortgage crisis. MBS involves three risks: interest rate risk, prepayment risk and default risk (Rosen 2007). Securitization of subprime mortgages increases the default risk. A recourse clause (in the representations and warranties) that obligates originators (lenders) to buy back loans that are later discovered not to have originated with proper due diligence became ineffective at reducing the risk because so many originators became insolvent. Securitization of mortgage was able to draw funds from the global financial market and fueled the housing bubble for an extended period. The globalization of mortgages made the mortgage market more opaque and the people with superior information took advantages of the less informed people. Therefore, those who had superior information increased their income at the cost of less informed people.

VI. Solutions to the Subprime Mortgage Crisis

VI.1 A Self-organization Perspective

One self-organization solution to the current financial crisis is to envisage the solution from the perspective of an individual financial institution, home buyer, regulation agency and investors in the financial system. A collapse of the housing bubble can be seen as the operation of negative feedback (Camazine, 2003), which offers an opportunity for error elimination (Plotkin, 1982; Popper, 1982). Individual financial institutions will react to the financial crisis based on their local information regarding the financial crisis. They are likely to scrutinize mortgage loans and to require a higher percentage of down payments. They will also interact with other actors in the system such as loan applicants, savers, other financial institutions and regulatory agencies, based on their local information. Financial institutions will exercise more due diligence in dealing with loan applicants and with mortgage-backed security issuers such as MBS and CDO issuers. They will pay special attention to CDSs. Investors will also scrutinize these financial instruments before they invest in them.

A general characteristic of a self-organizing system is robustness or resiliency (Camazine et al., 2003; Heylighen, 2008; Kauffman, 1993; Prigogine and Stengers, 1984). Self-organization in biology states that living organisms react to changes in environment, and their reactions lead to a global order (Camazine et al., 2003). The current perturbation will push financial systems into a better state or equilibrium, a state where individual financial institutions, consumers and regulators mutually adapt. Foster (1992, 2000) argues that the firm, as a complex adoptive system with self-organizational qualities, can develop a range of forward looking contractual arrangements in the context of transaction cost economics. The same can be said about the current changes in the financial environment. Financial firms will develop a range of new financial arrangements in dealing with the mispricing of risk and the opacity created by the originate and distribute mortgage financing model. Financial systems will establish a new global order stemming from the current perturbations. This self-organization perspective may provide an understanding the inside of the coordination process in “invisible hand” for Adam Smith. As Kauffman (1993) argues, “evolution is a complex combinatorial optimization process in each of the coevolving species in a linked ecosystem, where the landscape of each actor deforms as
the other actors move” (p. 644). Financial institutions securitized subprime mortgages and made mortgage loans available to low income home buyers, which made home buyers buy more homes. The result deformed the final holders of CDOs and led to the current crisis. The crisis, a stimulus to the financial system and a complex combinatorial optimization process in each of the coevolving home buyers, financial institutions, final investors (holders) and regulatory agencies, will help create new systems. The mutual interactions of these actors help generate more healthy and efficient financial intermediaries. Kauffman (1993) also points out that evolution is an emergent order honored and honed by selection. Emergence of a new financial system likewise will have the order honed by selection, and the new order will show how order emerges from the chaos. Moreover, the situation can become worse before it gets better, and the U.S. financial systems will go through the rugged landscape until it reaches an equilibrium or attractor. The experience on the 1997 Korean economic crisis provides an evidence of the self-organization perspective solution. However, alternative solutions may be required to reduce pain in the short-run and to trigger economic recovery.

VI.2 Alternative Solutions to Self-organization
Alternative solutions to self-organization perspectives include central authorities, blueprints, templates and leadership. The U.S. government and the Federal Reserve are central authorities; they are working to solve problems created by the housing bubble formed by positive feedback of self-organization. The collapse of the housing bubble has created enormous adverse effects on the U.S. and world economies. This negative feedback is a step in the process of reaching a new equilibrium from the collapse of the housing bubble. Camazine et al. (2003) point out that the “individual acquires and processes information that elicits a negative feedback response: A small perturbation applied to the system triggers an opposing response that counteracts the perturbation” (p. 16). In a biological system, this negative feedback prevents an implosion. An ideal solution to problems in the housing bubble might have been due diligence exercised by the regulatory agencies while the housing bubble was forming and expanding. Can regulatory agencies play this role? This may be a challenging task, because non-linear dynamic self-organization models are capable of generating catastrophic discontinuities, chaotic dynamics and a variety of other complex dynamics, as noted by Rosser (2000). Home buyers and financial institutions were interacting with each other based on local information. Home buyers responded to increasing housing prices, and financial institutions responded to opportunities for fee generation from mortgage origination, servicing mortgages, issuing MBSs and CDOs and underwriting mortgages and MBSs. They did not see problems stemming from the global pattern, group characteristics and complexity. Regulatory agencies thought that actors in markets are smart and control themselves to maintain balance. However, markets failed to self-regulate and negative feedback did not operate until the housing bubble reached a maximum. Therefore, a new regulatory regime may be required to prevent problems of bubbles from recurring. Today almost all large financial institutions in the U.S. and the world are involved in securitization as MBS and CDO issuers and underwriters. They actively participate in CDSs to manage their risk on CDOs. Consequently, problems faced by financial institutions, the U.S. economy and other economies in the world are severe and widespread. The crisis involved trillions of dollars worldwide, thus the problem requires massive coordinated efforts by governments in the world. To address this problem the U.S. Congress passed the Troubled Assets Recovery Program (TARP) and has designated $750 billion to purchase assets and equity from financial institutions. TARP, designed to strengthen the financial sector, allows the U.S. Department of Treasury to buy illiquid, toxic assets from banks and other financial institutions.
and to provide relief to homeowners who are facing mortgage bankruptcies. TARP also encourages banks to resume lending both to each other and to consumers and businesses. Interbank lending will restore financial market stability and make bank loans available to consumers and businesses, which will help increase consumer spending on durable goods such as automobiles, housing and furniture. However, toxic assets are likely to rise unless policy makers take measures to stop and reverse the trend. The U.S. government is taking steps to stem the foreclosure trend; it is planning to spend $75 billion for home owners who are facing home mortgage foreclosures and to offer subsidies to new home buyers to boost new home construction.

However, the costs of stabilizing financial systems are expected to rise significantly higher than the currently appropriated dollar amounts. U.S. policy makers may draw lessons from the Japanese real estate crisis in the 1990s. According to the New York Times (February 13, 2009: Hiroko Tabuchi, B 1), the Japanese economy endured a “lost decade” of economic stagnation as Japanese banks and policy makers were slow to recognize the magnitude of their banking problems and wasted trillions of Yen on half-measures. U.S. policy makers and banks need to confront the issues directly and should take effective steps fast. When the Korean government, banks and business firms took bold measures as they faced the 1997 financial crisis, they recovered relatively quickly (Park, 2008). Therefore, the timing and implementation of appropriate policy measures are crucially important in recovery of the U.S. economy from the current crisis. Current policy debates and institutional interventions on problems of self-organization have the same familiar tone of Keynesian and classical policy debates (Rosser, 1999). However, hindsight on the current financial crisis offers some clues for government interventions in the future. Since design of the originate and distribute model, the development of off-balance sheet special purpose vehicles (SPVs) and the securitization of subprime mortgages have contributed to the current financial crisis, these developments should have been properly monitored as they were developing. Are government agencies capable of doing the job? Scholars continue to debate this question.

The self-organizational policy prescription would further suggest that policy makers study the evolving nature of patterns (Colander, 2000) and address problems as they arise. Rather than bounding after the unknowable, and try to deduce analytically models that hold for all times, economics has reduced its search to what it believes is knowable. New Millennium economists search for patterns in data, try to find temporary models that fit the patterns, and study the changing nature of those patterns as institutions change. (Colander, 2000, p. 131)

Economists could not have known all potential problems when they designed and implemented the originate and distribute mortgage model. Problems emerged in the process of evolution in the new model. The securitization of subprime mortgage accompanied the cascade of events and a complex system has emerged as agents in financial institutions managed their credit risks and consumers acted on their local information of subprime mortgage securitization. The complexity of subprime mortgages led to catastrophic adverse effects on the global economy because the securitized subprime mortgages were sold to savers all over the world. Asymmetry of information between the originators of subprime mortgages and its final holders in other countries was more pronounced than other securities because of the complexity of subprime mortgage securitization. Therefore, government interventions require better understanding of the global patterns of bubbles (Rosser, 1999) and mechanisms for selecting a solution among tentative solutions (Popper, 1982) to reduce policy errors in mortgage financing changes. Policy
makers need to develop a process to mobilize knowledge from all knowledgeable people because people working in the field have concrete experiences and knowledge that the new system is experiencing and can help guide its future direction. Actions generate intended and unintended consequences (Giddens, 1984). Intended consequences of mortgage securitization are low mortgage rates, higher liquidity of mortgages and global market access to credit. Unintended consequences are special purpose vehicles (SPVs) and complex CDOs and CDSs. Unintended consequences necessitate further actions which can stem the adverse effects of actions. The U.S. financial market and government have failed to take steps to stem the adverse effects of the unintended consequences. Reasons for the failure of corrective actions or inactions are numerous, but the condition that prevailed at the time in the U.S. was the idea that the market knows the best. The current debt of the U.S. federal government limits fiscal policy options to address problems of the subprime mortgage crisis at this critical juncture. This illustrates that corrective action alternatives are conditioned on the situation of the time and space (now and in the U.S.). Making a correct policy choice requires consciousness and knowledgeable ability of actors (policy makers) and actors tend to have knowledge in the field. There were conscious voices among actors who were working in the field and became aware of the problems of subprime mortgages, but we ignored their voices. Therefore, it is necessary to establish channel knowledge of conscious actors to address problems of adverse consequences. We tend to pay more attention to the intended consequences of new policy, but often fail to have due diligence on the unintended consequences of new policy.

When the Federal Reserve and the new administration began to take more aggressive steps, the Fed made over one trillion dollars of loans to financial institutions in 2008, the Fed also purchased subprime mortgages and it is likely to maintain an easy monetary policy until early 2012. The new administration and Congress in 2009 worked out massive fiscal policy stimulus measures to boost the U.S. economy. The Bush administration had implemented tax cuts, but the amount of the cuts might not have been sufficiently large enough to trigger economic stimulation. The Obama administration had taken more aggressive stimulus fiscal policy measures. They proposed over $800 billion of recovery and reinvestment programs, and the U.S. Congress actually adopted the $787 billion stimulus and recovery package. Some economists argued that this amount was inadequate to address the problem of shortfalls in private sector spending. They argue that the spending gap was 1.3 trillion dollars.

Regulatory agencies are also scrutinizing financial institutions and their transactions more closely. Financial regulatory agencies are criticized for their lack of regulations and due diligence. The increasing tendency of deregulation and strong belief in a free market system in the past two decades created an environment for less stringent regulatory implementation. However, these trends are reversing now and some economists are proposing a financial product consumer protection act, equivalent to the current Consumer Protection Act. Opaqueness stemming from information asymmetry among actors in securitization of mortgage credit created moral hazard, adverse selection and conflict of interest. These problems may have contributed to the current subprime mortgage crisis. New or existing regulations need to reduce or eliminate the opacity of the system and increase transparency of the securitization process. The system should encourage the originator to take more risk rather than transferring it. Securitization of subprime mortgages and new issuances of CDOs and CDSs need to be monitored with due diligence.

Shiller (2008) has proposed that rather than more regulatory measures, we should have more democratization of financial systems to provide independent advice and information to financial consumers, so that consumers can make informed financial decisions. He also points out that the
collapse of the housing bubble makes housing more affordable to home buyers. Ultimately, it is important to design an efficient regulatory system, and the current subprime mortgage crisis presents an opportunity for decision makers to create efficient new orders for the U.S. and the global financial systems. The U.S. Congress actually passed a financial consumer protection act to protect financial consumers and to prevent recurring future financial problems. The new regulatory institution for monitoring financial market is the Consumer Finance Protection Bureau (CFPB).

VII. Summary and Conclusion
The originate-to-distribute model was originated to increase the liquidity of mortgages (lower mortgage rates) and to address interest rate risk stemming from lending long and borrowing short in mortgage financing (the problem of qualitative asset transformation). The new model facilitated the securitization of mortgages. However, it has created new problems of opacity, mispricing of risk and complexity in mortgage financing, as Popper (1982) predicted. Regulatory agencies need to understand these sources of opacity, mispricing of risk and complexity which developed underneath the housing bubble and securitization of mortgages to address ensuing problems. Long intermediation chains or complexity may help financial organizations and professionals, but hinder the supply of capital in real corporations. This issue needs to be carefully studied in future research. Securitization of subprime mortgages facilitated the growth of shadow banking which contributed to the current financial crisis. Therefore, policy makers have to be proactive in addressing problems that follow a new system because the new system continues to emerge as seen in the securitization of mortgages.

The pattern of the housing bubble or any bubble for that matter can be analyzed from a self-organization perspective. The current housing bubble was triggered by the securitization of subprime and Alt A mortgage loans by private financial institutions in 1995. The subprime mortgage crisis arose as the housing bubble popped, and the current crisis provides a good lesson why due diligence is required in new mechanism design (Maskin, 2008). Since no one would know all potential problems stemming from the emerging complexity of a new originate and distribute model at the time of its adoption, monitoring in a new model needs to be flexible and ongoing throughout the development of complexity in the model. Characteristics of complexity in self-organization are not universal and therefore controlling any bubble needs to be case-specific (Colander, 2000; Rosser, 1999). Although economic bubbles are recurring frequently and impose enormous adverse effects on millions of people as an economic bubble collapses, not enough studies have been conducted on economic bubbles. There is a need for more research to address problems effectively, and the self-organization perspective provides an appropriate framework for both the analysis and policy prescriptions of economic bubbles.

References
Foster, J., 2000, “Is there a role for transaction cost economics if we view firms as complex adaptive system?” Contemporary Economic Policy, 18(4), 369-385.
Haken, Herman, 1977, Synergestics, Berlin: Springer-Verlag Publishing Co.
Morgensen, 2008-02-17, New York Times, ISDA market survey.

Appendix A: Outstanding Mortgage Securities and Securitization Rate

<table>
<thead>
<tr>
<th>Year</th>
<th>GSE</th>
<th>Non-agency</th>
<th>Total Security</th>
<th>Total Mortgage</th>
<th>Securitization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>111,086</td>
<td></td>
<td>$111,086</td>
<td>$957,900</td>
<td>11.59%</td>
</tr>
<tr>
<td>1981</td>
<td>126,186</td>
<td></td>
<td>126,187</td>
<td>1,030,200</td>
<td>12.25</td>
</tr>
<tr>
<td>1982</td>
<td>162,829</td>
<td></td>
<td>162,892</td>
<td>1,070,200</td>
<td>15.22</td>
</tr>
<tr>
<td>1983</td>
<td>219,201</td>
<td></td>
<td>219,201</td>
<td>1,186,100</td>
<td>18.48</td>
</tr>
<tr>
<td>1984</td>
<td>252,007</td>
<td>11,000</td>
<td>263,007</td>
<td>1,132,100</td>
<td>19.91</td>
</tr>
<tr>
<td>1985</td>
<td>314,554</td>
<td>24,000</td>
<td>338,554</td>
<td>1,518,600</td>
<td>22.29</td>
</tr>
<tr>
<td>1986</td>
<td>434,884</td>
<td>16,600</td>
<td>451,484</td>
<td>1,722,000</td>
<td>26.22</td>
</tr>
<tr>
<td>1987</td>
<td>531,867</td>
<td>27,800</td>
<td>559,667</td>
<td>1,920,500</td>
<td>29.14</td>
</tr>
<tr>
<td>1988</td>
<td>570,733</td>
<td>34,900</td>
<td>605,633</td>
<td>2,154,100</td>
<td>28.12</td>
</tr>
<tr>
<td>1989</td>
<td>646,759</td>
<td>43,300</td>
<td>690,057</td>
<td>2,378,900</td>
<td>29.01</td>
</tr>
<tr>
<td>1990</td>
<td>1,013,920</td>
<td>55,000</td>
<td>1,068,920</td>
<td>2,614,700</td>
<td>40.88</td>
</tr>
<tr>
<td>1991</td>
<td>1,152,453</td>
<td>96,700</td>
<td>1,249,153</td>
<td>2,781,700</td>
<td>44.91</td>
</tr>
<tr>
<td></td>
<td>Prime</td>
<td>Subprime</td>
<td>Alt A</td>
<td>Total</td>
<td>[(2) + (3)]/(4)</td>
</tr>
<tr>
<td>---</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>1995</td>
<td>25,837.7</td>
<td>17,771.4</td>
<td>498.3</td>
<td>44,107.4</td>
<td>41.42%</td>
</tr>
<tr>
<td>1996</td>
<td>31,418.7</td>
<td>30,769.4</td>
<td>1,802.6</td>
<td>63,990.7</td>
<td>50.90</td>
</tr>
<tr>
<td>1997</td>
<td>49,974.9</td>
<td>56,920.7</td>
<td>6,518.0</td>
<td>113,413.7</td>
<td>55.93</td>
</tr>
<tr>
<td>1998</td>
<td>97,365.2</td>
<td>75,829.9</td>
<td>21,235.5</td>
<td>194,430.6</td>
<td>49.92</td>
</tr>
<tr>
<td>1999</td>
<td>74,630.9</td>
<td>55,851.5</td>
<td>12,022.8</td>
<td>142,505.2</td>
<td>47.63</td>
</tr>
<tr>
<td>2000</td>
<td>53,584.9</td>
<td>52,467.4</td>
<td>16,443.6</td>
<td>122,495.9</td>
<td>56.25</td>
</tr>
<tr>
<td>2001</td>
<td>142,202.5</td>
<td>87,052.9</td>
<td>11,373.6</td>
<td>240,629.0</td>
<td>40.90</td>
</tr>
<tr>
<td>2002</td>
<td>171,534.4</td>
<td>122,680.9</td>
<td>53,462.7</td>
<td>347,780.8</td>
<td>50.66</td>
</tr>
<tr>
<td>2003</td>
<td>237,454.6</td>
<td>194,958.5</td>
<td>74,151.0</td>
<td>506,564.1</td>
<td>53.12</td>
</tr>
<tr>
<td>2004</td>
<td>233,378.1</td>
<td>362,549.3</td>
<td>158,585.8</td>
<td>754,513.2</td>
<td>69.07</td>
</tr>
<tr>
<td>2005</td>
<td>280,703.7</td>
<td>465,036.3</td>
<td>332,323.2</td>
<td>1,078,063.2</td>
<td>73.96</td>
</tr>
<tr>
<td>2006</td>
<td>219,037.4</td>
<td>448,599.6</td>
<td>365,675.8</td>
<td>1,033,312.8</td>
<td>78.80</td>
</tr>
<tr>
<td>2007</td>
<td>180,462.4</td>
<td>201,546.7</td>
<td>249,610.0</td>
<td>631,619.1</td>
<td>71.42</td>
</tr>
</tbody>
</table>

Sources: Columns (1), (2), (3), (4) are from the 2008 Mortgage Finance Market Statistical Annual, Vol. I, II (Inside Mortgage Finance). Column (5) is calculated by the author. Columns (1), (2), (3) and (4) are billions of dollars.
Commercial Banks in Investment Banking Underwriting: Certification Effect or Conflicts of Interest

Yong-Chern Su
Department of Finance
National Taiwan University,
50 Lane 144 Sec. 4, Keelung Road, Taipei, Taiwan

Yu-Shin Lin
Department of Finance
National Taiwan University
50 Lane 144 Sec. 4, Keelung Road, Taipei, Taiwan

Han-Ching Huang*
Department of Finance
Chung Yuan Christian University
Business Building 404A, 4/F, 200, Chung Pei Road
Chung Li, Taiwan
samprass@cycu.edu.tw

* Corresponding author
Abstract

We focused on the case that China Development Industrial Bank (CDIB) bid for Grand Cathay Securities Corporation (GCSC) by the tender offer in 2001. Because of variance effect, arbitrage chance, and uncertainty problems of CDIB, tender offer announcement had a positive effect on GCSC’s stock returns but a negative one on CDIB’s stock returns. Owing to signaling effect, expectation of control right concentration, and increasing difficulty of taking control for CDIB, management resistance announcement had a positive effect on GCSC’s stock returns but a negative one on CDIB’s stock returns. Because of arbitrage chance, expectations of management resistance, and uncertainty problems of CDIB, tender offer approval by SFC had a positive effect on GCSC’s stock returns but a negative one on CDIB’s stock returns. Owing to disappointment to the result, tender offer failure had a negative effect on GCSC’s stock returns and a negative one on CDIB’s stock returns.

JEL classification: G01; G14; G21

Keywords: Tender Offer; Strategies; Event Study; Abnormal Returns; Cumulative Abnormal Returns

Introduction

In the last two decades, the tender offer was very popular in U.S. There were also many literatures about it (e.g. Louis, et al., 2010; Rustige and Grote, 2011). From the traditional issues including rationales of mergers and acquisitions (M&A) and their effect on corporation securities value to the studies which specifically focused on characteristics of the hostile takeover target, bid premium determinants in a tender offer and the effect of the tender offer on stock returns of both target and acquirer. However, there were much little studies on the tender offer in Taiwan. There may be some reasons for it. First, because of incomplete regulations of tender offer, insufficient incentives and shortage of human resources of M&A in the past, little companies adopted the tender offer as a way to take control of the target companies. Second, because of little companies, which adopted the tender offer, there were few samples that researchers could study. Therefore, in Taiwan, we lack complete and sufficient studies on the tender offer, especially in the field of finance.

However, with the development of M&A in Taiwan, the requirements of studies on the tender offer become more and more important. Especially in the financial industry, with the need of competition after entering WTO, the development of laws concerned (such as “Law Governing Merger in Financial Institutions” and “Financial Holding Company Laws”) and the encouraging attitude towards financial reform of the R.O.C. Government, to establish a complete legal system of the tender offer is essential. Additionally, it is also essential to study why companies should adopt the tender offer (how to give them incentives), how they can implement the process (how to decrease the difficulties to implement) and what affects a tender offer will cause.

Therefore, we choose the case that China Development Industrial Bank (CDIB) bid for Grand Cathay Securities Corporation (GCSC) by means of a tender offer in 2001 (the first case of tender offer in Taiwan’s financial industry and the amount of this bidding reached about NT$2.8 billion), and we determine our purposes in this article as follows. First, analyze why CDIB desired to bid for a securities corporation. Second, analyze why GCSC became a hostile takeover target. Third, analyze why CDIB adopted the tender offer as a way to bid for GCSC. Fourth, describe the process of tender offer CDIB implemented and the result of this case. Fifth, examine the effect of a tender offer on stockholders’ wealth of both CDIB and GCSC.
The rest of this paper is organized as follows. In Section 1, we introduce the data and explain the techniques used to measure the effect of a tender offer on stock returns of CDIB and GCSC. In Section 2, there are three topics we analyze. First, we describe the process and result of this case and the after development of relationship between CDIB and GCSC in 2002. Second, we analyze the strategic decision of CDIB on why it desired to acquire a securities corporation, why the target was GCSC and why CDIB chose tender offer as a tool to take control of GCSC. Third, we examine the effects of important events in the process of this case on the stock returns of both CDIB and GCSC. In Section 3, we summarize our findings and suggest some directions for future research.

I. Data and Methodology

A. Data

Our event days are defined by means of identifying the first announcement day of every event reported in the Economic Daily News from March 28, 2001 through April 8, 2002. The two firms in the case of tender offer we study are acquirer, China Development Industrial Bank (CDIB) and target firm, Grand Cathay Securities Corporation (GCSC). We acquire the daily stock closing quotes, dividends, and daily adjusted prices of CDIB listed on Taiwan Stock Exchange (TSE) market and those of GCSC listed on Over-The-Counter (OTC) Securities Exchange market from Taiwan Economic Journal (TEJ) Data Bank during the period from February 1, 2000 to April 10, 2002. The daily returns on the market portfolio are defined as the returns on TSE Weighted Index and OTC Securities Exchange Index at the ending of each day. We also acquire daily closing quotes of TSE Weighted Index and those of OTC Securities Exchange Index from TEJ Data Bank during the period starting on February 1, 2000 and ending on April 10, 2002.

B. Methodology

We define five events including tender offer announcement, management resistance, tender offer approval by SFC, tender offer failure and subsequent success of taking control to test the effects of these events on both stock returns of China Development Industrial Bank (CDIB) and those of Grand Cathay Securities Corporation (GCSC). And we define the event periods based on not calendar days but transaction days, for example, a day –2 relative to event day does not mean that two calendar days before the event day but two transaction days before it.

(1) The Event of Tender Offer Announcement

Because CDIB generated information to stock market that it would purchase shares of GCSC through a tender offer on March 28, 2001, we define that March 28, 2001 is the event day to test the tender offer announcement effect on both stock returns of acquirer and that of target firm. And then we define the event day as ”0” and the period starting on day –2 and ending at day +2 relative to event day as event period of tender offer announcement effect.

(2) The Event of Management Resistance

Because the board of GCSC made an announcement that it would resist the offer by CDIB on April 2, 2001, we define that April 2, 2001 is the event day to test management resistance announcement effect on both stock returns of acquirer and that of target firm. And then we define the event day as ”0” and the period starting on day -1 and ending on day +1 relative to event day as event period of management resistance announcement effect.
Because the board of GCSC approved of reducing the seats of board of directors from 11 to 9 on April 30, 2001, we define that April 30, 2001 is the event day to test anti-takeover measures effect on both stock returns of acquirer and that of target firm. And then we define the event day as “0” and the period starting on day -1 and ending on day +1 relative to event day as event period of anti-takeover measures effect.

(3) The Event of Tender Offer Approval
This case of tender offer is approved by SFC on April 19, 2001, we define April 19, 2001 is the event day to test tender offer approval effect on both stock returns of acquirer and that of target firm. And then we define the event day as “0” and the period starting at day -1 and ending at day +1 relative to event day as event period of tender offer approval effect.

(4) The Event of Tender Offer Failure
On May 25, 2001, in the stockholder’s meeting of GCSC, incumbents acquired 5 directors in the board, which contained 11 directors, but CDIB only acquired 2 directors. Therefore, we define May 25, 2001 is the event day to test tender offer failure effect on both stock returns of acquirer and that of target firm. And then we define the event day as “0” and the period starting at day -1 and ending at day +1 relative to event day as event period of tender offer failure effect.

(5) The Event of Subsequent Success of Taking Control
Although CDIB lost in the contest of GCSC’s control in the case of the tender offer, its dream for acquiring GCSC’s control came true on April 8, 2002. On this day, CDIB announced that it had entered into an agreement with GCSC’s major stockholders (Kuomintang (KMT)-related investment corporations) to purchase about 0.53 billion shares of GCSC from them. We define that April 8, 2002 is the event day to test the success of taking control effect on both stock returns of acquirer and that of target firm. And then we define the event day as “0” and the period starting on day -2 and ending at day +2 relative to event day as event period of success of taking control effect.

The followings define the estimation periods.
(1) The Event of Tender Offer Announcement
We define the event day March 28, 2001 as “0” and the period starting on day -302 and ending on day -3 relative to event day as the estimation period to test tender offer announcement effect.

(2) The Event of Management Resistance
We define the event day April 2, 2001 as “0” and the period starting on day -301 and ending on day -2 relative to event day as the estimation period to test management resistance announcement effect. We define the event day April 30, 2001 as “0” and the period starting on day -301 and ending on day -2 relative to event day as the estimation period to test anti-takeover measures effect.

(3) The Event of Tender Offer Approval
We define the event day April 19, 2001 as “0” and the period starting on day -301 and ending on day -2 relative to event day as estimation period to test tender offer approval effect.

(4) The Event of Tender Offer Failure
We define the event day May 25, 2001 as "0" and the period starting on day -301 and ending on day -2 relative to event day as the estimation period to test tender offer failure effect.

(5) The Event of Subsequent Success of Taking Control
We define the event day April 8, 2002 as "0" and the period starting on day -302 and ending on day -3 relative to event day as estimation period to test success of taking control effect.

We summarize the event days, event periods, and estimation periods mentioned as Table 1.

Abnormal Returns (AR) Measurements
(1) Defining Returns Measurements
We define stock $i$ daily returns at $t$ :

$$R_{it} = \ln \left\{ \left[ \frac{(P_{it} \times (1 + SD_{it} + NIR_{it}) + (CD_{it})]}{[(P_{it-1}) + (NIP_{it} \times NIR_{it})]} \right]\right\}$$

Where $P_{it}$($P_{t-1}$) : The closing quote of stock $i$ at $t$ ($t-1$)
$SD_{it}$ : The stock dividend rate per share of stock $i$ at ex-right $t$
$NIR_{it}$ : The offered rate per share in the capital increase by cash of stock $i$ at ex-right $t$
$NIP_{it}$ : The offered price per share in the capital increase by cash of stock $i$ at ex-right $t$
$CD_{it}$ : The cash dividend per share of stock $i$ at ex-dividend $t$

We define market daily returns at $t$ :

$$R_{mt} = \ln \left( \frac{P_{mt}}{P_{mt-1}} \right)$$

where $P_{mt}$($P_{m_{t-1}}$) : The closing market index at $t$ ($t-1$)

Additionally, because CDIB’s stocks and GCSC’s stocks are trading in the different securities exchange markets, we match CDIB’s stock daily returns with daily returns of TSE Weighted Index and match GCSC’s stock daily returns with daily returns of OTC Securities Exchange Index.

(2) Establishing Returns Model
We have tried the ARCH (1)-MA (1) model, GARCH (1,1)-MA (1) model and GARCH (1,2)-MA (1) model for returns model establishment. Eventually, we find that the ARCH (1)-MA (1) model is the best one, which can fit the data of daily stock returns of both CDIB and GCSC well. Therefore, we use ARCH(1)-MA(1) model as the returns model :

$$R_{it} = \alpha + MA_{it-1} + \beta * R_{mt} + \epsilon_{it}$$

$$\epsilon_{it} \sim N(0, h_{it})$$

$$h_{it} = A + B * \epsilon_{it-1}^2$$

where $R_{it}$ is stock $i$ daily returns at $t$. $R_{mt}$ is market daily returns at $t$

(3) Testing the Significance of Returns Model
a. Testing the Significance of Individual Coefficients
   This test is for whether the individual coefficient differs from zero. The test statistic we use to examine the significance of individual coefficients in ARCH model is t-statistic. We obtain the t-value of individual coefficients from
\[ t = \frac{\text{estimator} - \text{parameter}}{\text{estimated standard error of estimator}} \]  

(4)

b. Testing the Overall Significance of Returns Model

The testing method of overall model significance we use is the Likelihood Ratio Test (LRT). Let \( \theta \) be a vector of parameters to be estimated, and let \( H_0 \) specify some sort of restriction on these parameters. Let \( \theta_U \) be the maximum likelihood estimate of \( \theta \) obtained without regard to the constraints, and let \( \theta_R \) be the constrained maximum likelihood estimator. If \( L_U \) and \( L_R \) are the likelihood functions evaluated at these two estimators, then the likelihood ratio test statistic is

\[ \lambda = -2 \ln \left( \frac{L_R}{L_U} \right) \]  

(5)

Under regularity, the large sample distribution of \( \lambda \) is chi-squared, with degree of freedom equal to the number of restrictions imposed. In this study, the number of parameters is five.

(4) Abnormal Returns Measurements

We define daily abnormal returns (AR) of stock \( i \) at \( E \):

\[ \text{AR}_{iE} = R_{iE} - E(R_{iE}) \]  

(6)

Where \( R_{iE} \): The actual daily returns of stock \( i \) at \( E \) during event period

\( E(R_{iE}) \): The expected daily returns of stock \( i \) at \( E \) during event period

And

\[ E(R_{iE}) = \alpha + \beta * R_{mE} \]  

(7)

Where \( \alpha, \beta \) are parameters estimated from (3)

We define cumulative abnormal returns (CAR) of stock \( i \) during event period from \( t_1 \) to \( t_2 \):

\[ \text{CAR}_{i}(t_1, t_2) = \sum_{t=t_{1}}^{t_{2}} \text{AR}_{iE} \]  

(8)

where \( \text{AR}_{iE} \) is daily abnormal returns of stock \( i \) at \( E \).

II. Empirical Results

In this section, we use the event study approach and ARCH(1)-MA(1) model as a returns model to examine the effect of important events during the period of tender offer. First, take a look of stock price and trading volume of CDIB and GCSC during the period of tender offer (Figure 1 and Figure 2). We can find that CDIB’s stock price was falling until March 28, 2001 and the trading volume was declining after March 28, 2001. In contrast, GCSC’s trading volumes during the period of tender offer was obviously large than previous periods, and its stock price was more volatile during this period. After May 10, 2001, when the opened period in tender offer was over, the stock price of GCSC soon fell from NT$19.4 on May 10 to NT$14.8 on May 16. These data about the stock prices and trading volumes of CDIB and GCSC can be a foundation for our interpretation of empirical results.

First, from the point of view of synergy, CDIB’s decision to acquire GCSC, as mentioned before, will make revenue-enhancing, cost-reducing and financial synergy. However, it is difficult and uncertain for investors to evaluate the revenue-enhancing effect in the future and theoretically, revenue-enhancing effect itself is much hard to be specified due to much uncertainty and problems of cross selling. Additionally, because of little overlapped businesses between CDIB and GCSC (CDIB just desired to acquire GCSC for filling the shortage of its securities businesses), the cost-reducing effect is also little in this case. Therefore, the financial synergy comes from expanded assets and more balanced revenues structure should be clearer than operating synergy at that time. According to Galai and Masulis (1976) and Shastri (1982) arguments, the variance effect of the merger will exist and affect the individual companies’
bonds value. In their model, there are two companies in the merger, and they argued that the company whose variance of cash flows is larger than the other one will have a positive variance effect on its bonds value in the merger; in contrast, the company whose variance of cash flows is smaller than the other one will have a negative variance effect on its bonds value in the merger. In the case which we study, CDIB’s revenues structure is more riskful than GCSC because most of its revenues don’t come from fixed fees but capital gains. Therefore, we can reasonably assume that the variance of cash flows of CDIB should be larger than that of GCSC. According to the argument of variance effect, there should be a negative effect on GCSC’s bonds value but a positive effect on CDIB’s bonds value. Under the assumption that the value of a company is fixed, there should be a positive effect on GCSC’s stocks value but a negative effect on CDIB’s stocks value, which corresponds to the empirical result we show.

Figure 1 CDIB Stock Price and Trading Volumes During the Period of Tender Offer

Source: Taiwan Economic Journal Data Bank

Figure 2 GCSC Stock Price and Trading Volumes During the Period of Tender Offer

A. The Announcement Effect on the Stock Prices of both Acquirer and Target

Source: Taiwan Economic Journal Data Bank

Our estimates of returns model during estimation period of event of tender offer announcement are shown as Table 1.

CDIB made the information about its intention to acquire GCSC by a tender offer to flow to the stock market on March 28, 2001. We can find that GCSC’s stockholders had significantly high positive cumulative abnormal returns, which reached about 26 percent. However, during the event period, CDIB’s stockholders had slight negative cumulative abnormal returns, which reached −6 percent (Table 2). We interpret this result as follows.
Table 1 Summary of Event Days, Event Periods, and Estimation Periods Defined

<table>
<thead>
<tr>
<th>Events</th>
<th>Event Days</th>
<th>Event Periods*</th>
<th>Estimation Periods**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tender Offer Announcement</td>
<td>March 28, 2001</td>
<td>-2 ~ +2</td>
<td>-302 ~ -3</td>
</tr>
<tr>
<td>Management Resistance Announcement</td>
<td>April 2, 2001</td>
<td>-1 ~ +1</td>
<td>-301 ~ -2</td>
</tr>
<tr>
<td>The Seats of Board of Directors are Reduced</td>
<td>April 30, 2001</td>
<td>-1 ~ +1</td>
<td>-301 ~ -2</td>
</tr>
<tr>
<td>Tender Offer Approval</td>
<td>April 19, 2001</td>
<td>-1 ~ +1</td>
<td>-301 ~ -2</td>
</tr>
<tr>
<td>Tender Offer Failure</td>
<td>May 25, 2001</td>
<td>-1 ~ +1</td>
<td>-301 ~ -2</td>
</tr>
<tr>
<td>Subsequent Success of Taking Control</td>
<td>April 8, 2002</td>
<td>-2~+2</td>
<td>-302~3</td>
</tr>
</tbody>
</table>

* relative to event day as "0", ** relative to event day as "0"

Table 2 Tender Offer Announcement-Abnormal Returns (AR) and Cumulative Abnormal Returns (CAR)

<table>
<thead>
<tr>
<th>Firm</th>
<th>Abnormal Returns (AR) During Event Period</th>
<th>Cumulative Abnormal Returns (CAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1</td>
<td>-2</td>
</tr>
<tr>
<td>GCSC</td>
<td>0.029</td>
<td>0.071</td>
</tr>
<tr>
<td>CDIB</td>
<td>0.006</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Second, from the point of view of expectations in the market, investors will expect that there may be a chance for arbitrage because of tender offer premium, as a result, to buy in GCSC’s stock (we can find that the trading volumes of GCSC’s stock significantly bursted in rising during the event period, even a little before the event period), which also causes GCSC’s stock price to rise. For CDIB, however, it is still doubtful whether CDIB can acquire GCSC
successfully and the content of an offer had not been approved yet at that time. Furthermore, CDIB was also involved in the problem of inside trading at that time (CDIB let information leak out before application). All of these have a negative effect on CDIB’s stock price.

B. The Resistance Effect on the Stock Prices of both Acquirer and Target

There were two events about GCSC’s management resistance happening in April, 2001. The first event was that on April 2, GCSC’s board announced their intentions to resist the offer CDIB provided and the second one was the plan proposed by GCSC’s board on April 30, 2001 to reduce the seats of GCSC’s directors.

We find that in both of events GCSC’s stockholders had positive cumulative abnormal returns but CDIB’s had negative ones. In the first event, GCSC’s stockholders had about 15.2% cumulative abnormal returns and CDIB’s had -2.2% ones (Table 3). In the second event, GCSC’s stock cumulative abnormal returns was 10.9% lower than those in the first event and CDIB’s stock cumulative abnormal returns was -5.5% also lower than those in the first one (Table 4).

Table 3 Resistance Announcement -Abnormal Returns (AR) and Cumulative Abnormal Returns (CAR)

<table>
<thead>
<tr>
<th>Firm</th>
<th>Abnormal Returns (AR) During Event Period</th>
<th>Cumulative Abnormal Returns (CAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>GCSC</td>
<td>0.071</td>
<td>0.044</td>
</tr>
<tr>
<td>CDIB</td>
<td>-0.006</td>
<td>-0.015</td>
</tr>
</tbody>
</table>

Table 4 Announcement of Director Seats Reduced-Abnormal Returns (AR) and Cumulative Abnormal Returns (CAR)

<table>
<thead>
<tr>
<th>Firm</th>
<th>Abnormal Returns During Event Period</th>
<th>Cumulative Abnormal Returns (CAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>GCSC</td>
<td>0.040</td>
<td>0.051</td>
</tr>
<tr>
<td>CDIB</td>
<td>-0.003</td>
<td>-0.017</td>
</tr>
</tbody>
</table>

There are two reasons for GCSC’s positive cumulative abnormal returns in events of resistance. The first reason is that management resistance may be a positive signal about the target firm value (Baron, 1983; Bagnoli and Lipman, 1989). Because target firm management has private information about the true value of the firm, their decision to resistance shows that retaining the firm is more valuable than just tendering their stocks. The second one is that management resistance announcement or defensive measures may show that board of target will try to concentrate the control right or repurchase the stocks outstanding. It provides positive expectation about target firm’s stock price. And as mentioned proceeding, GCSC had abnormal brokerage and trading volume on its stocks in the open market so may strengthen the positive expectation of its stock price.

For the acquirer-CDIB, GCSC’s resistance will increase the difficulty to take control of the target. It implies that CDIB may need to offer much higher premium to target’s stockholders, spend more resources or time to seek for other block holder’s supports and so on. Therefore, GCSC’s resistance causes the negative effect on CDIB’s stock returns.

Additionally, the reason why GCSC’s positive cumulative abnormal returns in the first event was higher than those in the second one may be that the first announcement of resistance had more
dominant effect or brought more shocks to the investors and it provided more expectations and imaginations about stock price due to no specific resistance form described. For CDIB, the reason why its negative cumulative abnormal returns in the second event is lower than those in the first one may be that the shares CDIB acquired in the tender offer were certain not to be helpful for taking control in 2001, so the proposal made by GCSC’s board to reduce seats of directors caused CDIB’s intention to take control of GCSC more difficult to realize and the probability of tender offer success much lower. Hence, the second event had more negative effect on CDIB’s stock returns.

C. The Approval Effect on the Stock Prices of both Acquirer and Target
CDIB’s application for a tender offer was approved by SFC on April 19, 2001. We can find that GCSC’s stock had positive cumulative abnormal returns, which reached 0.9 percent, but CDIB’s stock had negative ones, which reached –1.4 percent (Table 5). The reasons of this result are similar to that of the first event-tender offer announcement. There is a motivation for investors to buy in GCSC’s stock for arbitrage. For example, we find that the GCSC’s stock price was only NT$16.7 on April 19 (the offer price of CDIB was NT$20), but quickly rose at the price of NT$19.03 on April 23 after approval. The trading volumes of GCSC stocks also increased obviously on following days, especially on April 20 and April 23. Additionally, after the approval of tender offer, investors also expected the management of GCSC would repurchase stocks from the market further, which also motivated investors to buy in GCSC’s stock.

Table 5 Tender Offer Approval-Abnormal Returns (AR) and Cumulative Abnormal Returns (CAR)

<table>
<thead>
<tr>
<th>Firm</th>
<th>Abnormal Returns During Event Period</th>
<th>Cumulative Abnormal Returns (CAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>GCSC</td>
<td>-0.017</td>
<td>-0.038</td>
</tr>
<tr>
<td>CDIB</td>
<td>-0.006</td>
<td>0.003</td>
</tr>
</tbody>
</table>

For CDIB, there was still high uncertainty whether the case of tender offer would succeed. Additionally, because the stocks which CDIB acquired by the tender offer didn’t have voting rights of stockholders’ meeting in 2001, it was doubtful whether the premium CDIB offered was worthy. Therefore, CDIB’s stock price didn’t perform well during the event period.

D. The Takeover Failure Effect on the Stock Prices of both Acquirer and Target
In the GCSC’s stockholders’ meeting on May 25, 2001, Jen-Hwa Investment Holding Co. acquired 5 directors and one auditor, however, CDIB only acquired two directors. Therefore, in 2001 CDIB could not obtained the control right and the intention of tender offer failed. During the event period, we can find that GCSC’s stock had slightly negative cumulative abnormal returns, which reached 0.2 percent, and CDIB’s stock also had negative ones, which reached –1.9 percent (Table 6). As mentioned before, there is a positive effect on GCSC’s stock returns but a negative effect on CDIB’s stock returns during the event period of tender offer announcement. When a tender offer failed, there should be a reverse effect on the stock returns of both GCSC and CDIB. We can find that the result of GCSC’s stock returns corresponded to the inference above, however, that of CDIB’s stock returns didn’t correspond to the inference above. The reason why this result happened may be that the investors were disappointed in the result of a tender offer, especially after CDIB’s previous large amount of investments in premiums offered. CDIB spent about NT$2.8 billion to purchase stocks tendered, but still lost the contest of control taking. It means that CDIB not only couldn’t enjoy any benefits of the merger coming from synergy but also lost a large amount of money in premiums offered.
Therefore, there might be a negative effect on CDIB’s stock returns during the event period of tender offer failure.

Table 6 Tender Offer Failure-Abnormal Returns (AR) and Cumulative Abnormal Returns (CAR)

<table>
<thead>
<tr>
<th>Firm</th>
<th>Abnormal Returns During Event Period</th>
<th>Cumulative Abnormal Returns (CAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>GCSC</td>
<td>-0.019</td>
<td>0.012</td>
</tr>
<tr>
<td>CDIB</td>
<td>-0.003</td>
<td>-0.015</td>
</tr>
</tbody>
</table>

E. Subsequent Stock Acquisitions Announcement Effect on the Stock Prices of both Acquirer and Target

After one year, the intention of CDIB to take control of GCSC came true. On April 8, 2002, CDIB announced that it would acquire GCSC’s stocks from Jen-Hwa Investment Holding Co. which agreed to this deal, and CDIB would hold about 0.74 billion shares, which was about 62 percentage of GCSC’s stocks. Principally, the acquiring price which CDIB planned wouldn’t exceed NT$21 per share. After this deal, CDIB would establish a foundation for GCSC’s being affiliated with the group of China Development Financial Holding Company (CDFHC). During the event period, we can find that GCSC’s stock had positive cumulative abnormal returns, which reached 3.29 percent, and CDIB’s stock also had positive ones, which reached 4.06 percent (Table 7). This result seems to show that investors had a positive attitude towards the success of this stock acquisition, and thought of that CDIB could take control of GCSC as good news. According to previous interpretation in the event of tender offer announcement, the result of success stock acquisition should be a positive effect on GCSC’s stock price and negative effect on CDIB’s stock price. However, we think there are some reasons about this contradiction. First, with many financial holding companies setting-up, the speed of financial service expanding and internal integration become a very important factor for whether a financial holding company can achieve competitive advantage, even survives in the future. Therefore, investors not only observe whether a financial institution becomes a financial holding company but also observe whether a financial holding company has a plan to strengthen its competence in the future. At that time, the group of CDFHC only included CDIB and First Taiwan Securities Corporation (FTSC) which was just a small-size securities corporation. Therefore, CDIB could acquire GCSC’s major block of stocks really made the market become more confident of CDIB’s future performance. Second, because of less uncertainty CDIB faced in this case compared with previous case of tender offer, there was not as many negative effects on CDIB’s stock price as that in the tender offer process one year before. Third, because this case was not a hostile takeover any more, there would not be anti-resistance costs expected to be spent. In sum, these may be the reasons why there was a positive effect on CDIB’s stock prices during event period, which was different with the result in the event of tender offer announcement one year before.

Table 7 Subsequent Stock Acquisition Announcement -Abnormal Returns (AR) and Cumulative Abnormal Returns (CAR)

<table>
<thead>
<tr>
<th>Firm</th>
<th>Abnormal Returns (AR) During Event Period</th>
<th>Cumulative Abnormal Returns (CAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>GCSC</td>
<td>0.0074</td>
<td>0.0061</td>
</tr>
<tr>
<td>CDIB</td>
<td>0.0129</td>
<td>0.0021</td>
</tr>
</tbody>
</table>
III. Conclusions

In this article, we focused on the case that China Development Industrial Bank (CDIB) bid for Grand Cathay Securities Corporation (GCSC) by means of the tender offer in 2001. First, we analyzed the reasons why CDIB desired to bid for a securities corporation. We found that this decision was based on responding to requirements of future competition, the encouraging attitude of the government towards M&A of financial institutions, strengthening CDIB’s securities businesses, low cost-of-entry of securities industry at that time and generating synergy. However, in the analysis of sources of synergy, we thought that there was little cost-reducing synergy but much possibly coming from revenue-enhancing and financial synergy. Additionally, we also analyzed why CDIB wanted to choose stock acquisition as a way to enter securities industry. The reasons included legal limitations on internal development, too risky and too slow to set up a subsidiary company, sufficient free cash flow CDIB owned and laying a foundation for changing into a financial holding company in the future.

Second, we thought the reasons why GCSC became a target of hostile takeover included its advantage on the businesses of securities underwriting and bonds trading, complete securities businesses, declining financial performance, huge capital base and stable financial structure, not being controlled by a certain family and convenience.

Third, we thought the reasons why CDIB adopted the tender offer as a tool for taking control of GCSC included opposed attitude of GCSC’s block holders, the failure of previous plan of open market sweep and the characteristics of ownership structure of GCSC.

Finally, we examined the effect of important events on stock returns of both CDIB and GCSC. We found that tender offer announcement had a positive effect on GCSC’s stock returns but a negative one on CDIB’s stock returns; management resistance announcement had a positive effect on GCSC’s stock returns but a negative one on CDIB’s stock returns; the decision GCSC’s board made to propose to reduce the seats of board had a positive effect on GCSC’s stock returns but a negative one on CDIB’s stock returns; tender offer approval by SFC had a positive effect on GCSC’s stock returns but a negative one on CDIB’s stock returns; tender offer failure had a negative effect on GCSC’s stock returns and a negative one on CDIB’s stock returns. Additionally, subsequent success of CDIB’s taking control of GCSC one year later had a positive effect on GCSC’s stock returns and a positive one on CDIB’s stock returns, too.

Eventually, we have some suggestions about the future directions for further studies. First, the regulations of tender offer and other laws concerned should be further studied for improvement, especially in reducing the difficulties of implementing a tender offer, rules of information disclosure and preventing inside trading problems. Second, as the samples in Taiwan increase in the future, the effect of important events of a tender offer, such as announcement or management resistance, on stock returns of both acquirer and target can be further and completely examined. Third, the characteristics of target in the cases of tender offers in Taiwan can be further studied and examined in the future.

References


The Determinants of Capital Structure in A Tax-Free Environment: An Extreme Bounds Analysis

Imad Moosa
Monash University and Kuwait University

Wafaa Sbeti*
American University of Kuwait
Liberal Arts Building
13034 Safat, Kuwait
wsbeiti@auk.edu.kw

*corresponding author
Abstract

Extreme bounds analysis is used to identify the determinants of capital structure in a tax-free environment, using data from Oman. The results show some evidence for the importance of liquidity, tangibility, growth opportunities and stock price performance. The findings are more supportive of the pecking order theory than the trade-off theory, but in general they imply that the capital structure decision does matter, even in the absence of taxes.

Keywords: Extreme Bounds Analysis, Capital Structure, Modigliani-Miller Theorem

I. Introduction

The choice of capital structure is one of the most important topics in corporate finance. The seminal work of Modigliani and Miller (1958), which is the cornerstone of capital structure theory, has led to the conclusion that financial leverage does not affect the firm’s market value. However, the Modigliani-Miller analysis was based on a set of highly restrictive assumptions that do not hold in the real world (perfect capital markets, homogeneous expectations, and the absence of taxes and transaction costs). Modigliani and Miller argued that, under these strict conditions, a firm’s value depends solely on the level and risk of its future cash flows. If this is the case, firms will be indifferent between financing operations by resorting to internal funds or by using different forms of external funds. This implies that an optimal capital structure does not exist because a firm’s value cannot be affected by the financing mode (debt versus equity).

The Modigliani-Miller work has motivated a vast amount of literature, leading to the formulation of alternative theories, such as the static trade-off model and the pecking order theory. These theories suggest a number of specific factors that may affect the capital structure. For example, the trade-off theory postulates that an optimal capital structure does exist and that a firm sets a target debt level and gradually moves towards it. The firm’s optimal capital structure involves a trade-off among the effects of corporate and personal taxes, bankruptcy costs and agency costs. In contrast, the pecking order hypothesis proposed by Myers and Majluf (1984) rejects the idea of a well-defined target debt ratio. According to this theory, retained earnings are the main source of funds for investment opportunities, followed by less risky debt, and last comes risky external equity financing. This is so because of the existence of an asymmetric information problem between inside and outside investors.

The presence of taxes provides an important reason for firms to seek an optimal capital structure, given that interest payments on debt are tax deductible. Typically, studies of capital structure have been conducted in the presence of taxes. It may therefore be interesting to examine the determinants of capital structure in a tax-free environment. For this purpose we use data from a tax-free country, Oman. To examine the determinants of capital structure in a tax-free environment, we use the technique of extreme bounds analysis (EBA), as suggested by Leamer (1983, 1985), and the extensions proposed by Granger and Uhlig (1991) and by Sala-i-Martin (1997). By using this technique, we test the robustness of coefficient estimates to changes in the conditioning set of information (represented by the explanatory variables). The advantage of EBA is that it makes it possible to avoid a problem that often arises in cross-sectional studies—that of the choice of the combinations of explanatory variables in the “optimal model”—and also the problem of data mining and biases in the reporting of results (estimating one thousand regressions and reporting the one that we fancy). We will return to this point later on.

II. Theories of Capital Structure

The conclusion of Modigliani and Miller on the irrelevance of the capital structure to the value of
the firm is inconsistent with real-world observations, where capital structure matters and firms are extremely reluctant to finance a project with a hundred per cent debt. As a result, financial economists have been on a quest to come up with the conditions under which the capital structure would indeed matter, as well as the factors that determine the choice between debt financing and equity financing. Broadly speaking, this strand of research has produced three theoretical approaches to the identification of the firm-specific factors that affect a firm’s capital structure. The three theories are discussed in turn.

The Trade-Off Theory
According to static trade-off model, a firm sets a target debt level and moves towards it gradually. As its name implies, the static trade-off theory explains observed capital structures in terms of a static trade-off between the costs and benefits of debt. It is the oldest theory and is directly related to the Modigliani-Miller (1958) irrelevance doctrine. The theory postulates that firms raise their debt level to the extent that the marginal tax advantages of additional borrowing are offset by the increase in the cost of financial bankruptcy. This is because raising more debt leads to tax benefits, but an increase in debt boosts the probability of default and hence the expected cost of bankruptcy.

Jensen and Meckling (1976) state that a value-maximizing firm pursues an optimal capital structure by considering the marginal costs and benefits of additional units of financing, and then choosing the form of financing that equates these marginal costs and benefits. Benefits of debt include the tax advantage and reduced agency costs of free cash flows. Costs include higher increasing risk of financial distress and the higher monitoring and contracting costs associated with higher debt levels. Thus, the trade-off theory of capital structure suggests that a firm’s target leverage is driven by three factors: taxes, the cost of financial distress, and agency conflicts. The trade-off theory postulates that since less profitable firms provide low shareholder returns, greater leverage in these firms merely boosts bankruptcy risk and the cost of borrowing, hence reducing return on shareholders’ equity still further, which in turn limits the potential for equity issue. Therefore, firms with low profits facing positive-NPV investment projects avoid external financing in general and leverage in particular, while markets will be reluctant to provide capital to such firms. The trade-off theory predicts a positive relation between leverage and profitability, and also between leverage and tangible assets.

The Pecking Order Theory
The pecking order theory was pioneered by Myers and Majluf (1984). This theory (based on information asymmetry) suggests that firms do not have a leverage target and that they focus on information costs and signaling effects. Myers and Majluf (1984) show that firms prefer to finance projects from internally generated cash flows—namely, retained earnings and depreciation expenses. When this source of funds is exhausted, they move on to debt. Additional equity is issued only when the latter is not sufficient to meet financing needs. This hierarchy is justified by differences in financing costs: issuing additional equity is the most expensive source of financing as it encompasses information asymmetries between managers, existing shareholders and potentially new shareholders. In view of the fixed payments associated with debt financing, it is less sensitive to information problems, while internally generated resources do not produce issuing costs.

In an attempt to explain this hierarchy, which is inconsistent with the prediction of the trade-off theory, Myers and Majluf (1984) constructed a model of information asymmetry assuming that
managers act on behalf of current shareholders. If external funds are needed to finance new investment, the market will interpret a new equity issue as indicating that the firm’s shares are overvalued, producing a negative impact on the share price. Thus, Myers and Majluf (1984) argue that if the firm does not have adequate funds to finance new investment, it will issue equity only when there are highly profitable investments that can be neither postponed nor financed through debt. Alternatively, this course of action will be followed when managers believe that the stock is so overvalued that shareholders will be prepared to tolerate a market-imposed penalty. Information asymmetry may cause current shareholders to renounce positive-NPV investment projects in order to avoid a drop in share price due to the issue of new equity, thereby causing an underinvestment problem. To avoid this outcome, firms implement financing policies that give them the capacity to finance investments while avoiding the use of external funds. As the pecking order theory predicts that firms use retained earnings first, then debt and equity issues as a last resort, firms with low profits facing positive-NPV investment projects will be more willing to use external funds if cash flows are weak. Therefore, there is a negative relation between leverage and profitability. Myers (1984) suggests that issuing debt secured by collateral may reduce the costs of financing related to asymmetric information. Difference in the information sets available to the various parties involved may lead to moral hazard (hidden action) and/or adverse selection (hidden information). Hence, debt secured by collateral may mitigate the costs of financing related asymmetric information, implying a positive relation between tangibility and financial leverage.

The Agency Cost Theory
Jensen and Meckling (1976) define agency costs as the sum of the monitoring expenditure of the principal, the bonding costs incurred by the agent, and a residual loss. The agency cost theory states that an optimal capital structure is determined by minimizing the costs arising from conflicts between the parties involved. Jensen and Meckling (1976) argue that agency costs play an important role in financing decisions due to the conflict that may exist between shareholders and creditors. If a firm is approaching financial distress, shareholders may encourage the management to take decisions that effectively expropriate funds from creditors to shareholders. Creditors will then require higher return for their funds if there is potential for this transfer of wealth. Debt and the accompanying interest payments, however, may reduce the agency conflict between shareholders and managers. Creditors have legal redress if the management fails to make interest payments when they are due. Hence, managers concerned about potential loss of job will be more likely to operate the firm as efficiently as possible in order to meet interest payments, thus aligning their behavior more closely with the objectives of shareholders’ wealth maximization.

Jensen and Meckling (1976) point out that debt agency costs arise due to a conflict of interest between debt providers (creditors), on one side, and shareholders and managers on the other side. Managers have an incentive to invest funds in risky business for shareholders’ interest because if the investment fails, creditors are likely to bear the cost (since shareholders have limited liability). The use of short-term debt may, however, mitigate the agency problems, as any attempt by shareholders to extract wealth from creditors is likely to restrict the firm’s access to short-term debt in the immediate future. Um (2001) suggests that if a firm’s level of tangible assets is low, the management (for reasons pertaining to monitoring costs) may choose a high level of debt to mitigate equity agency costs.

---

1 Agents incur bonding costs to reduce agency conflict.
Therefore, a negative relation between debt and tangibility is consistent with an explanation based on equity agency costs. He also argues that firm size may be a proxy for the debt agency cost (monitoring cost) arising from conflict between managers and investors. He asserts that the monitoring cost is lower for large firms than for small ones.

The Effect of External Factors
Apart from the firm’s own characteristics, external factors may affect capital structure. Gleason et al (2000) suggest that these factors include the legal environment, tax environment, economic system and technological capabilities. Likewise, Antoniou et al (2008) find that the capital structure decision of a firm is not only affected by its own characteristics, but also by the surrounding environment. The environment may affect capital structure for different reasons, such as deterioration or improvement in the state of the economy, the existence or otherwise of a stock market, and/or the size of the banking sector. Furthermore, Korajczyk and Levy (2003) argue that both macroeconomic conditions and firm-specific factors impinge upon financing choices. Deesomsak et al (2004) suggest that the capital structure is influenced by the environment in which they operate.

The Empirical Evidence
Prasad et al (2001) survey a large volume of literature on capital structure, concluding that the evidence on trade-off versus pecking order theories remains inconclusive. Booth et al (2001) argue that it is difficult to distinguish between trade-off and pecking order models because the variables used in one model are also relevant in the other model. The results provided by studies of the capital structure are a mixed bag. Shyam-Sunder and Myers (1999) test the static trade-off theory against the pecking order theory, using a sample of 157 U.S. firms over the period 1971 to 1989. They reject the trade-off model and find the pecking order model to have much greater explanatory power. Fama and French (2002) test the trade-off and pecking order predictions about debt and dividend for 3000 firms over the period 1965-1999. They find that each of the two models explains some of the financing behavior, and that neither of them can be rejected.


A limited number of empirical studies use data from developing countries. For example, Booth et al (2001) analyze data from ten developing countries (Brazil, Mexico, India, South Korea, Jordan, Malaysia, Pakistan, Thailand, Turkey and Zimbabwe) between 1980 and 1991. Pandy
(2001) uses data from Malaysia for the period 1984-1999, and Chen (2004) employs data from China for the period 1995-2000. Furthermore, Deesomsak et al (2004) investigate the determinants of capital structure in four Asia Pacific countries: Thailand, Singapore, Malaysia and Australia between 1993 and 2001. They suggest that the capital structure of firms operating in these countries is not only influenced by the determinants of capital structure that are widely defined in the literature, but also by the environment where they operate. Thus, while there is no agreement as to what determines capital structure decisions or if firms have an optimal capital structure, the environment where they operate affects their capital structures.

While most of the literature deals with the importance of firm-specific factors in determining a firm’s financing choice, the empirical evidence on the role played by stock market development is rather limited. For example, Demirguc-Kunt and Maksimovic (1996) conduct a comprehensive study that explores the effect of stock market development on firms’ financing choices. Using a sample of 30 industrial and developing countries for the period 1980-1991, they investigate the extent to which variation in aggregate debt-equity ratios within these countries can be explained by: (i) the level of development of the country’s financial markets; (ii) macroeconomic factors such as growth and inflation; and (iii) firm-specific factors that have been identified in the corporate finance literature as determining financial structure.

The results obtained by Demirguc-Kunt and Maksimovic (1996) reveal statistically significant negative correlation between stock market development (measured by the ratio of market capitalization to GDP) and the ratios of long-term and short-term debt to total equity. When they compare stock markets in developed and developing countries, they find that stock market development leads to a substitution of equity for debt financing in developing countries. In contrast, large firms in developed countries become more leveraged as the stock market develops, whereas small firms appear not to be significantly affected by market development. Their results have important implications for banks in emerging markets, since they do not need to be fearful of stock market development. They find that improvements in the functioning of a developing stock market result in a higher debt-equity ratio and thus more business for banks. In other words, stock markets and banks are complementary.

However, the Demirguc-Kunt and Maksimovic (1996) study suffers from a number of problems. In particular, it does not consider country-specific effects and uses data for both developed and developing countries from different sources, implying variations in the definition, collection procedure, and measurement of the underlying variables. Such problems render generalizations, comparisons and inferences difficult to make. Furthermore, their measures are averaged over the period 1980-1991, which is problematic, as many changes that occur simultaneously are ignored. Aggregation over time may blur important events and differences across countries.

Booth et al (2001) examine the capital structure decision in ten developing countries. They find the determinants of capital structure in these countries to be similar to those in developed countries. However, they reveal differences in the way leverage is affected by macroeconomic variables such as GDP growth, inflation and capital market development. Specifically, they find the relation between stock market activity and leverage to be significantly negative. However, they conclude that more research needs to be done in order to understand the impact of such factors on capital structure decisions.
Methodology

The methodology typically used in studies of capital structure is based on a regression of the form:

\[ L = \alpha_0 + \sum_{i=1}^{n} \alpha_i x_i + \epsilon \]  

(1)

where \( L \) is the leverage ratio and the \( x_i \)'s are explanatory variables. This regression is estimated by using cross-sectional or pooled data. Studies based on equation (1) invariably report a sample of regressions, including various combinations of the explanatory variables. The combinations are chosen for convenience because they produce “nice” results, in the sense that they vindicate the researcher’s pre-conceived ideas.\(^2\) The problem is that theory is not adequately explicit about what variables should appear in the “true” model.\(^3\) The following problem is often encountered: \( x_1 \) may be significant when the regression includes \( x_2 \) and \( x_3 \), but not when \( x_4 \) is included. So, which combination of all available \( x_i \)'s do we choose?\(^4\)

To test the importance of various explanatory variables in determining the capital structure while circumventing the problem of choosing the explanatory variables to appear in any regression, we employ extreme bounds analysis as developed by Leamer (1983, 1985) and extended by Granger and Uhlig (1990) and Sala-i-Martin (1997). This technique is used to find out if there is robustness in the determinants of the dependent variable. Hussain and Brookins (2001) argue that the usual practice of reporting a preferred model with its diagnostic tests, which is what was invariably done in previous studies, need not be sufficient to convey the degree of reliability of the determinants (the explanatory variables). However, EBA enables the investigator to find upper and lower bounds for the parameter of interest from all possible combinations of potential explanatory variables. The technique provides a useful method for assessing and reporting the sensitivity of the estimated coefficients to specification changes. Leamer and Leonard (1983, p 307) argue that the extreme values of the coefficient on the variable of interest delineate ambiguity in the inference about the coefficient induced by the ambiguity in choice of model—that is, model uncertainty. The relation between the dependent variable and a given explanatory variable is considered to be robust if the estimated coefficient on that variable remains statistically significant and correctly signed when the set of explanatory variables are changed.

EBA is based on a linear regression that is used to explain the capital structure, proxied by the

\(^2\) Economists are notorious for estimating 1000 regressions, throwing 999 in the bin and reporting the one they like (the one that produces the results that the researcher likes). While true scientific research should be based on a quest for the “truth”, it is unfortunate that the endeavour is usually for proving a pre-conceived idea. We often come across statements like “…unfortunately, the results turned out to be disappointing”. Gilbert (1986, p 288) casts significant doubt on the validity of the practice of assigning 999 regressions to the waste bin, because they do not produce the anticipated results. He further argues that it is this problem that has led Leamer (1983) to suggest that “econometricians confine themselves to publishing mappings from prior to posterior distributions rather than actually making statements about the economy”.

\(^3\) This would be the case if, for example, the final model specification is derived by solving a theoretical optimisation problem.

\(^4\) This is why those researchers reporting the “best” five or six regressions (which is quite common in the corporate finance literature) may find themselves in trouble. This happens when the five or six regressions give inconsistent results (\( x_i \) is significant in the first regression but not so in the fifth). In this case, the researcher may be tempted to conceal more of the truth by reporting the “one and only”, just one regression that tells a nice story.
leverage ratio, $L$, in terms of some explanatory variables. The model takes the form

$$L = \alpha + \sum_{i=1}^{n} \delta_i X_i + \beta Q + \sum_{i=1}^{m} \gamma_i Z_i + \varepsilon$$  \hspace{1cm} (2)$$

where $X_i$ is an explanatory variable that is always included in the regression because its importance has been established by previous studies, $Q$ is the variable whose robustness we want to test, and $Z_i$ is a potentially important variable. The $X_i$’s are called “free variables”, whereas $Q$ is called the “variable of interest”.\(^5\)

The technique centers around the estimated values of the coefficient on the variable of interest, $Q$. A large number of regressions are run to estimate the value of this coefficient, such that each regression contains the free variables, the variable of interest and a combination of a fixed number of $Z$ variables that are chosen from a predetermined pool. The procedure involves varying the set of $Z$ variables included in the regression to find the widest range of coefficients on the variable of interest, $\beta$, that standard hypothesis tests do not reject. By running a large number of regressions for each variable of interest, we identify the highest and lowest values of $\beta$ that cannot be rejected at a particular significance level. If the extreme values remain significant and of the same sign, then one can infer that the result (and hence, the variable of interest) is robust. Otherwise, the variable is fragile.\(^6\)

EBA has been criticized as being too stringent a test of robustness, in part because a variable is considered fragile even if one regression out of many thousands causes a change in the sign of a coefficient (the one-rotten-apple problem). Sala-i-Martin (1997) notes that if one keeps trying different combinations of control variables comprised of samples drawn with some error from the true population, then one is virtually guaranteed to find a model for which the coefficient of interest becomes insignificant or even changes sign. As a result, one may conclude either that no variables are robust or that the test of robustness is difficult to pass. Likewise, McAleer et al (1985) argue that without knowing the full set of characteristics of models generating extreme bounds, one cannot rely on EBA to test the robustness of any variable.

There have been a number of attempts to refine the robustness criteria in order to reduce the probability of obtaining unreasonable extreme bounds (for example, Granger and Uhlig, 1990). As a result, a reasonable EBA test has been developed to estimate the extreme bounds on the coefficient of interest by eliminating models with poor goodness of fit as measured by $R^2$. Granger and Uhling proposed this refinement of EBA by imposing a condition on the level of goodness of fit, such that all models with low $R^2$ are irrelevant for the calculation of extreme bounds. This criterion is represented by

$$R^2 \geq [(1 - \phi)R^2_{\text{max}} + \phi R^2_{\text{min}}]$$  \hspace{1cm} (3)$$

where $0 < \phi < 1$. If $\phi = 0$, the extreme bounds will be drawn from one model only, the one with

---

\(^5\) Free variables may also be chosen because their perceived relevance to the capital structure decision is highly intuitive or can be substantiated by strong theoretical reasoning. One free variable may be chosen because it outperforms the others in terms of these criteria.

\(^6\) A large number of regressions are required because of the large number of possible combinations of the $Z$ variables used with each variable of interest. Naturally, the number of regressions increases with the number of potential explanatory variables. For example, Sala-i-Martin (1997) ran almost two million regressions. In the original version of the paper, he ran about four million regressions (Sala-i-Martin, 1996).
the highest $R^2$, whereas if $\phi = 1$, then all models are relevant for the determination of the extreme bounds. Any other value means that extreme bounds are determined by models with an $R^2$ in the top $\phi$ per cent of the $(R^2_{\text{max}} - R^2_{\text{min}})$ range. This modification results in the so-called “restricted extreme bounds analysis” (REBA).

**The Choice and Measurement of Variables**

As we have said before, the existing literature does not offer a consistent theoretical framework for guiding empirical work on capital structure, because no single model specifies a full list of the determining variables. This makes the choice of suitable explanatory variables potentially contentious (Titman and Wessels, 1988). As a result, we use a number of potential explanatory variables where robustness (with respect to their effect on capital structure) is determined by EBA. The variables used in the empirical analysis are listed in Table 1.

**The Dependent Variable: Leverage Ratio**

The literature does not provide a clear-cut definition of leverage, and the specific choice depends on the objective of the analysis. Rajan and Zingales (1995) apply four alternative definitions of leverage. The first and broadest definition is the ratio of total liabilities to total assets. This can be viewed as a proxy for what shareholders are entitled to in case of liquidation. However, this measure does not provide a good indication of whether or not the firm is exposed to the risk of default in the near future. In addition, since total liabilities include items like accounts payable, which are used for transaction purposes (as opposed to financing), it is likely to overstate the amount of leverage. Furthermore, this measure of leverage is potentially affected by provisions and reserves, such as pension liabilities.

The second definition of leverage is the ratio of debt (both short and long term) to total assets. This measure of leverage only covers debt in a narrow sense—that is, interest bearing debt, hence it excludes provisions. The measure also fails to incorporate the fact that some assets are offset by specific non-debt liabilities. For example, an increase in the gross amount of trade credit is reflected in a reduction in this measure of leverage. Because the level of accounts payable and accounts receivable may differ across industries, Rajan and Zingales (1995) suggest a measure of leverage that is unaffected by the gross level of trade credit.

The third definition is the ratio of total debt to net assets, where net assets are total assets less accounts payable and other current liabilities. This measure of leverage is unaffected by non-interest bearing debt and working capital management. However, it is influenced by factors that have nothing to do with financing. For example, assets held against pension liabilities may reduce this measure of leverage. The fourth and last definition is the ratio of total debt to capital, where capital is defined as total debt plus equity. This measure of leverage pertains to the capital employed and thus represents the effects of past financing decisions. It relates more directly to the agency problems associated with debt.

An additional issue is whether leverage should be computed as the ratio of the book or market value of equity. Fama and French (2000) argue that most of the theoretical predictions apply to book value. Likewise, Thies and Klock (1992) suggest that book ratios provide a better representation of the targets adopted by the management. The market value of equity depends on a number of factors that are out of the direct control of the firm. Furthermore, if only the market value of equity is used to compute leverage and firms use the book value of equity in their decisions on leverage, the leverage of firms whose market to book ratio is very high will be underestimated. Likewise, if only the book value of equity is used when measuring leverage and
firms use the market value of equity in their decisions on the level of debt financing, the level of leverage of firms that have high market to book ratio will be overestimated.

Another issue that pertains to the choice between book value and market value is the observation that book values tend to be influenced by the choice of accounting methods, whereas market values tend to vary considerably, which may result in changes in leverage without changes in either the amount of outstanding debt or the book value of equity. In this study, the choice falls on the ratio of the book value of total debt to the book value of total assets.

Explanatory Variable 1: Size (SIZ)
It is generally accepted that firm size is an important determinant of the ability of firms to raise capital through debt or equity. The majority of studies suggest a positive relation between leverage and size. The most important underlying argument is that information asymmetries are less severe for large firms than for small ones. If the public is more aware of what is going on at a large firm, the firm will find it easier to raise debt. Furthermore, large firms can diversify their investment projects on a broader basis and limit their exposure to cyclical fluctuations in one particular line of production. Thus, it is plausible to suggest that financial distress risk is lower for large firms.

The trade-off theory implies a positive relation between firm size and leverage, since large firms are exposed to lower bankruptcy risk and cost. In addition, large firms have lower agency costs of debt, relatively smaller monitoring costs, less volatile cash flows, and easier access to capital markets. They also require more debt to benefit fully from the tax shield. Titman and Wessels (1988) argue that large firms tend to be more diversified and fail less often, so size may be an inverse proxy for the probability of bankruptcy, which means a positive relation between size and the firm’s debt capacity. Also, large firms may be able to utilize the economies of scale in issuing long-term debt, and may even have bargaining power over creditors, thus implying lower borrowing costs. Many studies suggest that leverage increases with the value of the firm (Narayanan, 1988; Harris and Raviv, 1990; Stulz, 2000) and generally find leverage to be positively related to firm size (Marsh, 1982; Rajan and Zingales, 1995; Booth et al, 2001; Antoniou et al, 2008)

Following previous studies, we proxy firm size by the value of total assets. We expect size to be positively related to leverage for three reasons: (i) large firms tend to be more diversified, which means that they are less risky, and as consequence they have a lower probability of default; (ii) large firms may be able to reduce the transaction cost associated with debt; and (iii) information costs are lower for large firms because the quality of financial information improves and mistrust diminishes as the firm becomes larger.

Explanatory Variable 2: Liquidity (LIQ)
Liquidity is defined as the ratio of current assets to current liabilities. This ratio is a measure of the ability of the firm to cover its short-term financial commitments. The pecking order theory predicts that firms with high liquidity borrow less. In addition, managers can manipulate liquid assets in favour of shareholders, thus increasing the agency cost of debt. A negative relation between leverage and liquidity is expected, simply because using more debt means more liabilities, thus implying fewer current assets remaining after covering liabilities. Nevertheless, when firms employ more current assets, they can generate more internal funds that can be used to finance their investment activities. A negative relation between leverage and liquidity implies that firms finance their activities following the financing hierarchy of the pecking order theory.
Explanatory Variable 3: Profitability (PRF)

There are conflicting theoretical predictions on the effects of profitability on leverage. According to the pecking order theory, firms raise capital by resorting to retained earnings first, then to debt, and to issuing new equity as a last resort. The preference of retained earnings over external financing can be attributed to the costs associated with the presence of information asymmetries between managers and outside investors. Debt typically grows when investment exceeds retained earnings and declines when investment falls short of retained earnings. Accordingly, the pecking order theory predicts a negative relation between leverage and profitability. Jensen (1986) and Williamson (1988) describe debt as a “disciplinary device” that is used to ensure that managers pay out profits rather than build empires. Most of the empirical studies show that leverage is negatively related to profitability, which confirms the pecking order hypothesis. Friend and Lang (1988) and Titman and Wessels (1988) obtain supportive evidence for U.S. firms. Kester (1986) finds leverage to be negatively related to profitability in both the U.S. and Japan. Rajan and Zingales (1995), Booth et al (2001) and Antoniou et al (2008) confirm this finding.

In the trade-off theory, however, agency costs and bankruptcy costs push more profitable firms towards higher book leverage for at least three reasons: (i) expected bankruptcy costs decline when profitability rises; (ii) the deductibility of corporate interest payments induces more profitable firms to use debt financing; and (iii) in the agency models, higher leverage helps to control agency problems by forcing managers to pay more of the firm’s excess cash (Jensen and Meckling, 1976; Easterbrook, 1984; Jensen, 1986). Stronger commitment to the allocation of a larger fraction of their pre-interest earnings to debt payments suggests a positive relation between book leverage and profitability. In other words, profitable firms with excess cash flows need a high debt level to prevent managers from engaging in sub-optimal investment projects. Thus, a positive relation between profitability and leverage is expected.

Explanatory Variable 3: Asset Tangibility (TAN)

Titman and Wessels (1988), Rajan and Zingales (1995), and Fama and French (2000) argue that the ratio of fixed to total assets (tangibility) should be an important factor for leverage. According to the trade-off hypothesis, tangible assets act as collateral and provide security to creditors in the event of financial distress. Jensen and Meckling (1976) point out that the shareholders of leveraged firms tend to overinvest, which gives rise to the classical shareholder-bondholder conflict. However, if debt can be secured against assets, the borrower is restricted to using borrowed funds for specific projects. Creditors have an improved guarantee of repayment while the recovery rate is higher—that is, assets retain more value in the case of liquidation. The debt capacity should rise with the proportion of tangible assets, which means that the trade-off theory predicts a positive relation between leverage and the proportion of tangible assets.

Conversely, Grossman and Hart (1982) argue that the agency costs of managers consuming more than the optimal level of fringe benefits is higher for firms with lower levels of assets that can be used as collateral. Managers of highly leveraged firms are less able to consume excessive fringe benefits, since bondholders monitor such firms more closely. The monitoring costs of this agency relation are higher for firms with less collateralizable assets, which mean that these firms tend to choose higher debt levels to limit fringe benefits. This agency model predicts a negative relation between tangibility of assets and leverage. Jensen and Meckling (1976) point out that the agency cost of debt exists as the firm may shift to riskier investments after the issuance of debt,
thus transfer wealth from creditors to shareholders to exploit the option nature of equity. Since tangible assets can be used as collateral (thus reducing the creditor’s risk of suffering such agency costs of debt), a high fraction of tangible assets is expected to be associated with high leverage.

The majority of empirical studies find a positive relation between tangibility and leverage (for example, Marsh, 1982; Long and Malitz, 1985; Friend and Lang, 1988; Titman and Wessels, 1988; Harris and Raviv, 1990; Rajan and Zingales, 1995). Some empirical studies of developing countries find mixed results. For example, while Um (2001) reports a positive relation between tangibility and leverage for Korea, other studies find tangibility to be negatively related to leverage—for example, Booth et al (2001) for ten developing countries and Huang and Song (2002) for China. It is suggested, however, that this relation depends on the type of debt. Bevan and Danbolt (2002) also find a positive relation between tangibility and long-term debt, whereas a negative relation is observed for short-term debt and tangibility in the United Kingdom. Based on the available evidence, a positive or negative relation between tangible assets and leverage might be expected. We use the ratio of tangible assets to total assets to measure asset tangibility.

Explanatory Variable 5: Growth Opportunities (GOP)
The relation between the expected growth of a firm and its leverage ratio should be negative for two reasons. First, the cost of financial distress increases with expected growth, forcing managers to reduce debt in the capital structure (trade-off theory). Second, firms issue equity instead of debt when overvaluation leads to higher expected growth (information asymmetry). Sometimes the internal sources of funds may not be sufficient to finance investment opportunities, in which case external funds are needed. If firms require external financing, they issue debt before equity (the pecking order theory). Hence, growth opportunities should be positively associated with leverage (Kremp et al, 1999).

Jensen and Meckling (1976) and Myers (1984) argue that when firms issue debt, managers have an incentive to engage in asset substitution, thus transferring wealth away from bondholders to shareholders. It is envisaged that agency costs are higher for firms with substantial growth opportunities. Thus, the trade-off model predicts that firms with more investment opportunities have less leverage because of stronger incentives to avoid underinvestment and asset substitution that may arise from stockholder-bondholder agency conflicts. This prediction is reinforced by Jensen’s (1986) free cash flow theory, which predicts that firms with more investment opportunities have less need for the disciplinary effect of debt payments to control free cash flows. The trade-off theory predicts a negative relation between leverage and growth opportunities, since the market value grows at least in proportion to investment outlays.

Rajan and Zingales (1995), who use market to book ratio as a proxy for the level of growth opportunities available to the firm, suggest that one would expect a negative relation between growth opportunities and the leverage ratio. This is consistent with the proposition put forward by Jensen and Meckling (1976) and Myers (1977) who point out that a firm with a high level of growth opportunities can be expected to have a low leverage ratio. Thus, firms with large amounts of investment opportunities tend to have low leverage ratios. In addition, Myers (1999) suggests a negative relation between profitable investment opportunities and long-term debt. The cost of financial distress, which is associated with high leverage, may affect a firm’s ability to finance future growth. Therefore, Myers argues that the managers of firms with valuable growth opportunities should choose low leverage.

According to Lang et al (1996), leverage is negatively related to growth only for firms with low
Tobin’s Q ratio—that is, for firms whose growth opportunities are not recognised by the capital market. However, the negative relation between leverage and growth does not hold for firms or industries with high Tobin’s Q ratio. Following Myers (1984), Smith and Watts (1992), Rajan and Zingales (1995) and Lang et al (1996), we use Tobin’s Q as a proxy for firm’s growth opportunities. It is a rough measure of agency costs because it captures the changing relation between future investment opportunities and existing assets.7

Explanatory Variable 6: Payout Ratio (POR)
The payout ratio is likely to play an important role in the formulation of financing decisions, primarily due to market imperfections. Miller and Rock (1985) show that a firm’s dividend policy and financing policy are closely related. Boyle and Eckhod (1997) demonstrate that if capital gains taxes are higher than dividend income taxes, firms will pay out more and raise more capital. However, if higher dividends signal increased future earnings, the firm’s cost of equity will be lower, thus favoring equity to debt. This implies a negative relation between leverage and the payout ratio. The effect of dividend policy on capital structure is likely to be influenced by country-specific institutional factors that are beyond the control of the firm. The direction and significance of the relation between capital structure and payout policy should depend on the net impact of information asymmetries, agency costs, ownership structure, and the tax laws of the country where the firm operates. The payout ratio is measured by the ratio of dividends to net income.

Explanatory Variable 7: Share Price Performance (SPP)
According to the pecking order theory, information asymmetries between managers and outside investors force managers to sell equity at a discount. Managers offer such a discount when the benefit of raising external equity capital outweighs the cost implied by the discount. When shares are overvalued, a discount could be offered without any real loss in the wealth of existing shareholders. This is possible if equity is issued after an increase in the share price due to overvaluation. This suggests an inverse relation between share price performance and the leverage ratio. However, such an inverse relation may be observed due to statistical distortions as the market value of equity increases with the increase in share price, even without any further equity issue. Since the book leverage ratio is independent of this effect, the effects on these two measures taken together should reveal the cause and nature of the relation between leverage and changes in share price. In this study stock price performance is measured by the annual return on the stock (capital gains).

Explanatory Variable 8: Firm’s Age (AGE)
Older firms have longer track records and therefore a higher reputational value. Older firms should, ceteris paribus, have a higher target leverage ratio since they face a lower external financing premium on debt. Age is measured by the number of years since incorporation.

Explanatory Variable 9: Income Variability (VAR)
The trade-off theory suggests that higher income variability boosts the risk that a firm may not be able to cover its interest payments, leading to higher expected costs of financial distress. This

7 Tobin’s Q is measured as \( Q = \frac{(TA - EBV + EMV)}{TA} \), where \( TA \) is total assets, \( EBV \) is the book value of equity, and \( EMV \) is the market value of equity.
implies a negative relation between income variability and target leverage. At the same time, higher income variability makes Myer’s (1977) underinvestment problem less severe, thus reducing the agency costs of debt. If the latter effect dominates, a positive relation emerges between income variability and leverage (Cools, 1993, p 223). Income variability is measured as the standard deviation of the firm’s net operating income over an eight year period.

**Data and Model Estimation Results**

The data used in this study, obtained from the Gulf Investment Guide, covers 42 companies listed on the Stock Exchange of Oman. The implementation of extreme bounds analysis requires each regression to contain a set of free variables that are always included in the model, the variable of interest and the Z variables. For theoretical and empirical considerations, only size is used as a free variable, as the importance of this variable has been established. In addition to the free variable and the variable of interest, three Z variables from a remaining list of seven are included in each regression.

This exercise involves running a total of 245 regressions, 35 regressions for each of the variables of interest with three Z variables and one X variables. In Figure 1 we see that some estimated coefficients change sign (and even significance), hence the underlying variables cannot be robust (GOP, SPP, AGE and VAR). Out of the remaining four variables, PRF seems the closest thing to being a robust variable, which is confirmed by the results presented in Table 2. \( \beta_{\text{max}} \) and \( \beta_{\text{min}} \) are reported, together with their t statistics and the percentage of regressions producing significant \( \beta \)’s at the 5 per cent significance level. These are, therefore, the results of traditional EBA. They show that the only robust variable is profitability. Notice that other variables would appear to be significant if the reported regressions are selected in such a way as to show “nice” results. These variables include liquidity, tangibility, growth opportunities and stock price performance.

Table 3, on the other hand, reports the results of restricted EBA, where \( \beta_{\text{max}} \) and \( \beta_{\text{min}} \) are picked from the equations with the highest 40 per cent \( R^2 \). The results show that there are three robust variables: liquidity (negative), profitability (negative) and stock price performance (positive). These results are more supportive of the pecking order theory.

**Concluding Remarks**

In this paper we used extreme bounds analysis to identify the determinants of capital structure of a cross-sectional sample of firms from Oman, where they operate in a tax-free environment. In the absence of a specific theoretical model with defined explanatory variables, EBA is used to circumvent the problem of choosing combinations of explanatory variables, which typically arises in cross-sectional studies, particularly in corporate finance.

Since EBA is a stringent test of robustness, only a few of the variables suggested in the corporate finance literature as determinants of the capital structure appear to be robust. Evidence is obtained for the robustness of liquidity, profitability and stock price performance. Judged by the signs of the estimated coefficients, the results are more supportive of the pecking order theory.

---

8 The *Gulf Investment Guide* is a periodical issued by Zugaibi and Kabbani, financial consultants based in Jeddah, Saudi Arabia. It covers shareholding companies of the six member countries of the Gulf Co-operation Council.

9 There is no specific rule as to the number of Z variables to be included in each regression. For example, Hussain and Brookins (2001) use four variables, while Levine and Renelt (1992) use three variables only to avoid multicollinearity. For a given pool of Z variables the smaller the number of variables chosen to be included in the equation the larger the number of equations that have to be estimated.
that the trade-off theory. Since there is evidence for the importance of some factors in determining capital structure, it is plausible to conclude that the capital structure decision does matter, even in the absence of taxes.

References


Table 1: A List of Variables and Their Definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leverage Ratio</td>
<td>LEV</td>
<td>Ratio of book of debt to book value of assets</td>
</tr>
<tr>
<td>Size</td>
<td>SIZ</td>
<td>Total assets</td>
</tr>
<tr>
<td>Liquidity</td>
<td>LIQ</td>
<td>Ratio of current assets to current liabilities</td>
</tr>
<tr>
<td>Profitability</td>
<td>PRF</td>
<td>Return on assets</td>
</tr>
<tr>
<td>Asset Tangibility</td>
<td>TAN</td>
<td>Ratio of tangible assets to total assets</td>
</tr>
<tr>
<td>Growth Opportunities</td>
<td>GOP</td>
<td>Tobin’s Q</td>
</tr>
<tr>
<td>Payout Ratio</td>
<td>POR</td>
<td>Ratio of dividends to net income</td>
</tr>
<tr>
<td>Share Price Performance</td>
<td>SPP</td>
<td>Changes in share price</td>
</tr>
<tr>
<td>Firm Age</td>
<td>AGE</td>
<td>Years since incorporation</td>
</tr>
<tr>
<td>Income Variability</td>
<td>VAR</td>
<td>Standard deviation of net operating income</td>
</tr>
</tbody>
</table>

Table 2: Results of Traditional EBA

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta_{max}$</th>
<th>t</th>
<th>$\beta_{min}$</th>
<th>t</th>
<th>Significant $\beta$’s at 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIQ</td>
<td>-0.022</td>
<td>-1.71</td>
<td>-0.077</td>
<td>-6.09</td>
<td>91.4</td>
</tr>
<tr>
<td>PRF*</td>
<td>-0.013*</td>
<td>-5.27</td>
<td>-0.038</td>
<td>-6.39</td>
<td>100.0</td>
</tr>
<tr>
<td>TAN</td>
<td>-0.152</td>
<td>-1.70</td>
<td>-0.500</td>
<td>-4.45</td>
<td>62.9</td>
</tr>
<tr>
<td>GOP</td>
<td>0.166*</td>
<td>2.68</td>
<td>-0.217*</td>
<td>-2.46</td>
<td>28.6</td>
</tr>
<tr>
<td>POR</td>
<td>-0.001</td>
<td>-1.81</td>
<td>-0.950</td>
<td>-1.25</td>
<td>31.4</td>
</tr>
<tr>
<td>SPP*</td>
<td>0.548*</td>
<td>3.01</td>
<td>-0.108</td>
<td>0.43</td>
<td>42.9</td>
</tr>
<tr>
<td>AGE</td>
<td>0.749</td>
<td>0.22</td>
<td>-0.926</td>
<td>-0.27</td>
<td>0.0</td>
</tr>
<tr>
<td>VAR</td>
<td>0.793</td>
<td>0.16</td>
<td>-9.86</td>
<td>-0.19</td>
<td>0.0</td>
</tr>
</tbody>
</table>

*Significant at the 5% level  
#Robust variable.

Table 3: Results of Restricted EBA (Regressions with the Highest 40% R-Squared)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta_{max}$</th>
<th>t</th>
<th>$\beta_{min}$</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIQ*</td>
<td>-0.029*</td>
<td>-2.76</td>
<td>-0.057*</td>
<td>-5.28</td>
</tr>
<tr>
<td>PRF*</td>
<td>-0.019*</td>
<td>-4.62</td>
<td>-0.037*</td>
<td>-6.02</td>
</tr>
<tr>
<td>TAN</td>
<td>-0.152</td>
<td>-1.70</td>
<td>-0.492*</td>
<td>-4.48</td>
</tr>
<tr>
<td>GOP</td>
<td>0.166</td>
<td>2.68</td>
<td>-0.042</td>
<td>-0.61</td>
</tr>
<tr>
<td>POR</td>
<td>-0.001*</td>
<td>-2.41</td>
<td>-0.950</td>
<td>-1.25</td>
</tr>
<tr>
<td>SPP*</td>
<td>0.530</td>
<td>3.65</td>
<td>0.417</td>
<td>2.36</td>
</tr>
<tr>
<td>AGE</td>
<td>0.749</td>
<td>0.22</td>
<td>-0.962</td>
<td>-0.27</td>
</tr>
<tr>
<td>VAR</td>
<td>0.591</td>
<td>0.16</td>
<td>-0.457</td>
<td>-0.11</td>
</tr>
</tbody>
</table>

*Significant at the 5% level  
#Robust variable.
Figure 1: Estimated Coefficients on the Variables of Interest (All Regressions)
Figure 1: Continued
Editorial Board

Editor-In-Chief
Prof. Y Rama Krishna, MBA, PhD

Associate Editor
Prof. Vinay Datar, Seattle University, USA

Senior Sub-Editor
Dr. Moustafa Abu El Fadl, Rhode Island College, USA

Members
Dr. Yu Hsing, Southeastern Louisiana University, USA
Prof. Richard J. Cebula, Jacksonville University, USA
Prof. Nur Adiana Hiau Abdullah, Universiti Utara Malaysia
Dr. Mohammad Talha, King Fahd University of Petroleum & Minerals, Saudi Arabia
Dr. Sorin A Tuluca, Fairleigh Dickinson University, USA
Dr. Nozar Hashemzadeh, Radford University, USA
Dr. Dar-Hsin Chen, National Taipei University, Taiwan
Dr. Hayette Gatfaoui, Rouen Business School, France
Dr. Emmanuel Fragnière, University of Bath, Switzerland
Dr. Shwu - Jane Shieh, National Cheng-Chi University, Taiwan
Dr. Raymond Li, The Hong Kong Polytechnic University, Hong Kong
Prof. Yang-Taek Lim, Hanyang University, Korea
Dr. Anson Wong, The Hong Kong Polytechnic University, Hong Kong
Dr. Shaio Yan Huang, National Chung Cheng University, Taiwan
Mr. Giuseppe Catenazzo, HEC-University of Geneva, Switzerland
Prof. J P Singh, Indian Institute of Technology (R), India
Dr. Jorge A. Chan-Lau, International Monetary Fund, Washington DC
Prof. Carlos Machado-Santos, UTAD University, Portugal
Dr. Osama D. Sweidan, University of Sharjah, UAE
Dr. David Wang, Chung Yuan Christian University, Taiwan
Prof. David McMillan, University of St Andrews, Scotland UK
Dr. Abdullah Sallehuddin, Multimedia University, Malaysia
Dr. Burak Darici, Balikesir University, Turkey
Dr. Mohamed Saidane, University of 7 November at Carthage, Tunisia
Dr. Anthony DiPrimio, Holy Family University, USA

Advisors
Dr. Michail Pazarskis, Technological Educational Institute of Serres, Greece
Dr. Huseyin Aktas, Celal Bayar University, Turkey
Dr. Panayiotis Tahinakis, University of Macedonia, Greece
Dr. Ahmet Faruk Aysan, Bogazici University, Turkey
Dr. Leire San Jose, Universidad del País Vasco, Spain
Dr. Arqam Al-Rabbaie, The Hashemite University, Jordan
Dr. Suleyman Degirmen, Mersin University, Turkey
Prof. Dr. Ali Argun Karacabey, Ankara University,
International Research Journal of Applied Finance
ISSN 2229 – 6891

Dr. Moawia Alghalith, UWI, St Augustine, West Indies
Dr. Michael P. Hanias, School of Technological Applications, Greece
Dr. Mohd Tahir Ismail, Universiti Sains Malaysia, Malaysia
Dr. Hai-Chin YU, Chung Yuan University, Taiwan
Prof. José Rigoberto Parada Daza, Universidad de Concepción, Chile
Dr. Ma Carmen Lozano, Universidad Politécnica de Cartagena, Spain
Dr. Federico Fuentes, Universidad Politécnica de Cartagena, Spain
Dr. Aristeidis Samitas, University of Aegean, Greece
Dr. Mostafa Kamal Hassan, University of Sharjah, College of Business, UAE
Prof. Muhammad Abdus Salam, State Bank of Pakistan, Pakistan