Asset pricing when trading is for entertainment

Jiang Luo
Nanyang Business School, Nanyang Technological University, Singapore, and
Avanidhar Subrahmanyam
Anderson Graduate School of Management,
University of California at Los Angeles, Los Angeles, California, USA

Abstract

Purpose – High levels of turnover in financial markets are consistent with the notion that trading, like gambling, yields direct utility to some agents. The purpose of this paper is to show that the presence of these agents attenuates covariance risk pricing and volatility, and implies a negative relation between volume and future returns. Since psychological literature indicates that the desirability of a gamble arises from the ex ante volatility of the outcome, the authors propose that agents derive greater utility from trading more volatile stocks. These stocks earn lower average returns in equilibrium, although the risk premium on the market portfolio is positive. The authors then consider a dynamic setting where agents’ utility from trading increases when they make positive profits in earlier rounds (e.g. due to an endowment effect). This leads to “bubbles,” i.e. disproportionate jumps in asset returns as a function of past prices, higher volume in up markets relative to down markets, as well as a leverage effect, wherein down markets are followed by higher volatility than up markets.

Design/methodology/approach – Analytical.

Findings – The presence of gamblers attenuates covariance risk pricing and volatility, and implies a negative relation between volume and future returns. If gamblers prefer more volatile stocks, these stocks earn lower average returns in equilibrium. If agents’ utility from trading increases when they make positive profits in earlier rounds (e.g. to an endowment effect), this leads to higher volume and lower volatility in up markets relative to down markets.

Originality/value – No paper has previously modeled agents who derive direct utility from trading.

Keywords Finance, Volume, Behavioural

Paper type Research paper

1. Introduction

The game of investing is intolerably boring and over-exacting to any one who is entirely exempt from the gambling instinct; whilst he who has it must pay to this propensity the appropriate toll. (Keynes, 1936)

Why do agents trade? In the neoclassical paradigm, trade occurs upon changes in market values of securities, or preferences or beliefs[1]. However, volume in financial markets appears to be too large to be explained by such considerations alone. Milgrom and Stokey (1982) imply that there should be no trade among investors with only speculative motives for trading. Barber and Odean (2000), however, document an average customer turnover of as high as 75 percent per year at a large discount brokerage firm. Annual turnover on the NYSE has ranged between 60 and 100 percent of shares outstanding over the past ten years, and is as high as about 80 percent in Germany and 150 percent in East Asian countries[2]. De Bondt and Thaler (1995, p. 392) note that “the high trading volume observed in financial markets is perhaps the single most embarrassing fact to the standard finance paradigm.” Noting the high volume in financial markets, Black (1986, p. 531) mentions the need to “introduce direct utility of trading” to explain trading activity.

Motivated by the above observations, we address the question: What is the nature of the equilibrium when agents directly obtain utility from the act of trading?[3] To the best of our
knowledge, an explicit analysis of this question, although suggested by Black (1986), does not yet appear in the literature. In our setting, information is symmetric and there are multiple assets, each traded by agents who possess a risk-averse utility function over wealth, but some of whom get additional utility from trading. The notion that agents may gamble for pleasure is well-established in literature (see, for example, Coventry and Brown, 1993 or Kuley and Jacobs, 1988). Indeed, the total revenue from casino gaming in the USA exceeded $70bn in 2015[4]. Work in neuro-psychology (Anselme and Robinson, 2007; Preuschoff et al., 2006) indicates that the *ex ante* uncertainty of the reward from gambling allows the secretion of dopamine, a pleasure-inducing hormone. Markiewicz and Weber (2013) show that gambling tendencies also create excessive stock trading[5]. These observations motivate our consideration of agents who participate in markets both for traditional wealth-related reasons and to obtain direct benefits from trading[6]. We show that our model and its extensions accord with a range of stylized facts, such as attenuation of covariance risk pricing (Fama and French, 1992), the low volatility anomaly (Ang et al., 2006), the volume premium in asset returns (Datar et al., 1998), higher volume in up relative to down markets (Karpoff, 1987), the leverage effect, wherein down markets accentuate volatility (Black, 1976; Christie, 1982), and provide additional untested implications.

Normal levels of trade to capture risk premia in markets are magnified when agents receive utility from the act of trading; this accords with volume levels greater than those expected from neoclassical models (Tkac, 1999). Interestingly, however, asset volatility is lower in the presence of such agents. This is because they act as *de facto* liquidity providers in equilibrium[7]. Specifically, conditional risk premia required to absorb liquidity shocks are lower when agents derive direct utility from trading. We also find that *β* pricing is obscured in equilibrium, and the degree to which it is obscured depends on the magnitude of trading utility. The intuition in our setting is that the compensation for risk demanded in equilibrium declines as this direct utility rises. So, expected returns are linear in *β*, but the coefficient on *β* attenuates as the direct utility from trading increases. This result accords with the notion that it is generally hard to find evidence of covariance risk pricing in equity markets (Haugen and Baker, 1996; Fama and French, 1992).

It is reasonable to suppose that agents may derive greater utility from some stocks relative to others. A central characteristic proposed by Kumar (2009) for stocks that are attractive to individual investors is volatility. Indeed, work (e.g. Linnet et al., 2012; Fiorillo et al., 2003) in psychology indicates that the pleasure from gambling emanates from the uncertainty or volatility associated with the outcome[8]. We thus consider the notion that agents obtain greater direct utility from trading stocks with more volatile payoffs. We find that under reasonable conditions, relatively volatile stocks get “overvalued” and earn negative future returns on a risk-adjusted basis. The intuition is that if volatile stocks yield greater consumption benefits from trading, this effect can more than counteract the tendency for risk aversion to induce covariance risk pricing. Since high payoff volatility, under reasonable conditions, also corresponds to high *β* and high idiosyncratic volatility (as we show), our analysis is consistent with the empirical result of Ang et al. (2006) that idiosyncratic volatility negatively forecasts asset returns. A corollary to our analysis is that high *β* (and high volatility) stocks are overvalued so that an investor should eschew them. Consistent with this notion, Frazzini et al. (2013, p. 4) indicate that a cornerstone of famed investor Warren Buffett’s investment philosophy is to buy “stocks that are ‘safe’ (with low *β* and low volatility).”

An interesting set of stylized facts in finance is that while volatility is negatively priced in the cross-section, on aggregate, the risk premium on equities is positive (Haugen and Baker, 2010; Mehra and Prescott, 1985). We show that our analysis accords with the “low volatility” anomaly wherein low risk stocks earn higher average returns than high risk stocks (as shown empirically in Baker and Haugen, 2012). At the same time, however, the
market portfolio commands a positive risk premium because our agents are risk averse. Thus, our analysis is simultaneously consistent with positive risk pricing in the aggregate, but negative pricing of volatility in the cross-section.

In our setting, when trading volume is high in a stock, it is associated with agents desiring greater (extra) utility from trading that stock. However, stocks from which agents derive greater utility from trading also tend to command higher prices and lower returns. Thus, our model is consistent with the negative cross-sectional relation between volume and returns documented in Datar et al. (1998). Additionally, we show that the relative attractiveness of alternative opportunities to participate in risky activities (such as casino gambling and lotteries) affects equilibrium trading volume in equities. Specifically, when alternative opportunities are more attractive, volume in equities is lower. This is consistent with Gao and Lin (2015), who show that equity trading volume in Taiwan decreases as the total jackpot of a major statewide lottery increases, indicating that stock trading acts as an alternative outlet for gambling beyond lotteries.

Previous literature argues that the pleasure from gambling rises when the outcome of a previous gamble is positive, either because of a “house money effect” or because of a positive physiological response to winning (e.g. Thaler and Johnson, 1990; Coventry and Constable, 1999; see also Barberis and Xiong, 2012). Accordingly, we consider an intertemporal setting where agents derive more utility from trading if previous rounds of trade have been profitable. We find that the impact of a positive piece of news can be greatly magnified. Specifically, as the news crosses a threshold, it causes a jump in the mass of agents who trade to derive direct utility, which results in an exaggerated response of asset prices to the news. This suggests “bubbles,” wherein later price moves represent overreactions to the initial price move, followed by subsequent corrections. Our finding accords with the experimental observation that excitement caused by up-moves in asset prices fuels bubbles (Andrade et al., 2016). Further, in down markets, there is less interest by agents who trade for entertainment, and thus lower liquidity provision, lower volume and higher volatility, which accords with the notion that up markets have higher volume than down markets (Karpoff, 1987; Chordia et al., 2007; Comiskey et al., 1987), and with the well-investigated leverage effect in financial markets (e.g. Black, 1976; Bekaert and Wu, 2000).

Our analysis suggests untested implications. Specifically, we argue that stocks in economies where retail investors (who are more likely to trade for entertainment) form a bigger fraction of the trading population will exhibit greater share turnover, and less evidence of covariance risk pricing. These stocks should also exhibit nonlinear responses to positive news. Within an economy, greater retail holdings predict lower pricing of covariance risk.

The idea that agents may trade for purposes of deriving enjoyment from trading has been discussed in the empirical literature. For example, Dorn and Sengmueller (2009) show, using survey data matched with actual trades, that agents who state that they derive “enjoyment” from trading turn over their portfolios to a greater degree than other investors. In a survey of retail investors, Dhar and Goetzmann (2006) state that more than 25 percent of investors view stock market investing as a hobby. Grinblatt and Keloharju (2009) provide empirical evidence that sensation seeking personalities (e.g. those who get speeding citations) may obtain a thrill from the act of trading. Our work complements these papers by explicitly modeling agents for whom trading yields consumption benefits in addition to influencing wealth. Note that such agents do not trade on mistaken beliefs and thus are not isomorphic to irrational agents. Indeed, they simply maximize a different objective relative to “traditional” investors who maximize the expected utility of wealth alone.

This paper is organized as follows. Section 2 presents the basic model, while Section 3 compares economies with and without agents who receive direct utility from trading. Section 4 endogenizes investors’ participation in the stock market to gamble. Section 5
presents a dynamic extension, and Section 6 concludes. All proofs of propositions and corollaries, unless otherwise stated, appear in Appendix 1, while Appendices 2 and 3 present some ancillary derivations.

2. The model
There are two dates, 0 and 1, and \( N + K \) risky securities. At date 1, these securities pay liquidating dividends of \( V = (V_1, \ldots, V_N + K)' \), which follows a multivariate normal distribution. At date 0, investors trade these securities. The per capita supplies of these securities, \( S = (S_1, \ldots, S_N + K)' \), also follow a multivariate normal distribution with finite, bounded, variances. Security prices are indicated by \( P = (P_1, \ldots, P_N + K)' \) and are determined in equilibrium. There is also a riskless asset, the price and return of which are normalized to unity.

There are two groups of agents. First, there is a mass \( \rho \) of neoclassical utility-maximizing agents. Second, there is a mass \( 1 - \rho \) of agents we call “G traders”; these achieve direct utility from trading risky securities. One can view the former class of agents as sophisticated investors (possibly institutions), and at least part of the latter class as unsophisticated (e.g., individual) investors.

The \( i \)th trader is endowed with \( Q_i = (Q_{i1}, \ldots, Q_{iN + K})' \) units of risky securities and \( M_i \) units of the risk-free asset. We assume that \( Q_{ij} = s_{ij} \forall j = 1, \ldots, N + K \), so that the endowment comprises the entire supply of shares available for trade in the economy. His wealth levels at Dates 0 and 1 are, respectively, given by:

\[
W_{i0} = M_i + Q_i'P, \\
W_{i1} = W_{i0} + X_i'(V - P),
\]

where \( X_i = (X_{i1}, \ldots, X_{iN + K})' \) is the quantity of risky securities he holds after the trading is complete.

The utility function of the \( i \)th regular (non-G) trader is the standard exponential one:

\[
U(W_{i1}) = \exp(-\gamma W_{i1}),
\]

where \( \gamma > 0 \). Based on the normality assumption of our model, he chooses \( X_i \) to maximize:

\[
E[U(W_{i1})] = \exp[-\gamma W_{i0} - \gamma \frac{1}{2} X_i' E(V - P)^{-0.5} \gamma X_i \text{Var}(V) X_i].
\]

The first-order condition (foc) with respect to (wrt) \( X_i \) implies that his demand can be expressed as:

\[
X_{NG}(P) = \frac{1}{\gamma} \text{Var}(V)^{-1} (E(V) - P).
\]

However, the \( i \)th G trader has the utility function:

\[
U_G(W_{i1}, C_i) = \exp(-\gamma W_{i1} - C_i).
\]

The extra term \( C_i \) captures the direct utility from trading. We let \( C_i = 0.5 X_i' G X_i \), where \( X_i \) is the quantity of risky securities he holds after the trading is complete, and \( G \) is a diagonal, positive semi-definite matrix. Thus, the G trader’s utility function can be expressed as:

\[
U_G(W_{i1}, X_i) = \exp(-\gamma W_{i1} - 0.5 X_i' G X_i).
\]

This utility function captures the notion that the bigger the quantity of trade (in absolute terms), the bigger is the utility derived from trading. Thus, we model the notion that
increasing the scale of the transaction increases the “thrill” derived from this larger scale of “gambling” in financial markets (Dorn and Sengmueller, 2009[10]. Also, for the modeling of direct utility from trading, of course, the utility function can depend on the unsigned order via a variety of exponents. The specific convex (quadratic) parameterization we choose lends tractability and closed-form solutions without the need to resort to numerical methods. The G trader chooses $X_i$ to maximize:

$$E[U_C(W_{1i}, X_i)] = -\exp\left[-\gamma W_{i0} - \gamma \left[X'_i E(V-P) - 0.5\gamma X'_i \text{Var}(V) X_i + 0.5X'_iGX_i/\gamma\right]\right].$$

The foc wrt $X_i$ implies that his demand can be expressed as:

$$X_C(P) = \frac{1}{\gamma} \left(\text{Var}(V) - G/\gamma^2\right)^{-1} (E(V) - P).$$

The soc holds under the assumption that $\text{Var}(V) - G/\gamma^2$ is a positive definite matrix.

As can be seen, the “numerator” in the above demand is the same as for non-G traders. However the position is larger per unit gain relative to non-G traders, i.e. the G traders take more aggressive positions relative to traditional utility maximizers. The parameters in the $G$ matrix apply only to risky assets, which is consistent with the notion that it is gambling on uncertain outcomes that yields direct utility. For now, we assume that $G$ does not vary with the riskiness of the assets, and relax this assumption in Section 2.5. Note that as the utility of trading increases, the position vector explodes, and beyond a certain level of $G$, there is no interior optimum. The scale of the position taken per unit expected price appreciation increases in $G$, which governs how much additional utility is derived from trading.

### 2.1 Risky security payoffs—the factor structure

We now explicitly model security payoffs as a factor structure to analyze how volume, volatility and the pricing of covariance risk are affected by the presence of G traders. In what follows, unless otherwise specified, a generic random variable, $\tilde{\eta}$, follows a normal distribution with mean zero and variance $\nu_\eta$.

The payoff of the $j$th ($j = 1, \ldots, N$) risky security takes a factor expression:

$$V_j = \tilde{V}_j + \sum_{k=1}^K \left(\beta_{jk}\tilde{f}_k\right) + \tilde{\epsilon}_j. \quad (2)$$

All $\tilde{f}$’s and $\tilde{\epsilon}$’s follow independent normal distributions. The payoff of the $N + K$ ($k = 1, \ldots, K$) risky security equals the realization of $k$th factor, i.e., $V_{N+k} = \tilde{f}_k$. Thus, this risky security can be a portfolio mimicking the $k$th factor.

For tractability, we use the risky securities to construct portfolios mimicking the $K$ factors and $N$ residuals. We refer to these portfolios as the basic securities. Use $\tilde{\theta}_j$, $j = 1, \ldots, N + K$, to denote the payoffs of the basic securities. Specifically, the first $N$ $\tilde{\theta}$’s indicate the payoffs of the $N$ residuals, that is, $\tilde{\theta}_j = \tilde{\epsilon}_j$ for $j = 1, \ldots, N$. The next $K$ $\tilde{\theta}$’s indicate the payoffs of the $K$ factors, that is, $\tilde{\theta}_{N+k} = \tilde{f}_k$ for $k = 1, \ldots, K$. These portfolios are orthogonal to each other, simplifying the exposition[11].

The per capita supplies of the $j$th basic security is indicated by $\tilde{\zeta}_j + \tilde{\epsilon}_j$, where $\tilde{\zeta}_j$ is a constant. Each agent’s endowment in a basic security equals the per capita supply of that security. Henceforth, we will assume that the mean supply of securities is positive (i.e. $\tilde{\zeta}_j > 0 \ \forall j$). The supply noise is not necessary for most of our main results (except those on price volatility). To simplify our analysis, we assume that the variance $\nu_\zeta,
is sufficiently small. Specifically, we assume that:

\[ \nu_i < \frac{1}{\gamma^2 \nu_0} \min(1/4, \rho/2). \]  

(3)

This condition facilitates the derivation of the results because it ensures that the \( G \) traders’ penchant for aggressively “buying low and selling high,” which drives many of our results, is not too adversely affected by excessive supply noise.

We denote the utility from trading the \( j \)th basic security as \( G_j \). Now, note that \( \tilde{\theta}_j \sim N(0, \nu_0) \). Each non-\( G \) and \( G \) trader’s demands for the \( j \)th basic security are given by:

\[ X_{NGj}(P_j) = \frac{E(\tilde{\theta}_j) - P_j}{\gamma \text{Var}(\tilde{\theta}_j)} = -\frac{P_j}{\gamma \nu_0}, \]

(4)

\[ X_{Gj}(P_j) = \frac{E(\tilde{\theta}_j) - P_j}{\gamma (\text{Var}(\tilde{\theta}_j) - G_j/\gamma^2)} = -\frac{P_j}{\gamma (\nu_0 - G_j/\gamma^2)}. \]

(5)

The denominator in Equation (5) needs to be positive for the second-order condition of the \( G \) traders to be satisfied. The intuition is that the effect of \( G_j \) alone is to induce \( G \) traders to want to trade infinite quantities. This tendency, however, is tempered by risk aversion and the volatility of the asset’s final payoff. Henceforth, we will assume that the condition for an interior optimum is indeed satisfied, i.e. that \( G_j < \gamma^2 \nu_0 \).

The modeling of direct utility from trading is distinct from other ways to generate volume such as differences of opinion (e.g. Varian, 1985; Kandel and Pearson, 1995). In our setting, agents have the same beliefs about the asset’s final payoff, yet there is trading. Indeed, it can be seen from Equation (5) that as \( G \) approaches its upper limit, the trade of \( G \) traders becomes infinitely large. Thus (as we will see in Section 2.4), there is high volume, even though agents agree on the asset’s payoff. There also is a distinction between volume induced by risk-seeking behavior and that induced by our \( G \) traders. With pure risk-seeking behavior, the positions of agents in risky assets are unbounded, precluding an equilibrium. In our model, an equilibrium is possible because \( G \) trading is tempered by risk aversion.

The market clearing condition requires:

\[ \bar{c}_j + \tilde{z}_j = \rho X_{NGj}(P_j) + (1-\rho)X_{Gj}(P_j), \]

from which we derive the prices and returns (i.e. price changes)[12] as presented in the following proposition:

**P1.** The price and return of the \( j \)th basic security are given by:

\[ P_j = -\gamma a_j(G_j) (\bar{c}_j + \tilde{z}_j), \]

\[ \tilde{R}_j = \tilde{\theta}_j - P_j = \tilde{\theta}_j + \gamma a_j(G_j) (\bar{c}_j + \tilde{z}_j), \]

where:

\[ a_j(G_j) = \frac{1}{\nu_0 + \frac{1-\rho}{\nu_0 - G_j/\gamma^2}}. \]

The term \( a_j(G_j) \) represents the effect of \( G \) traders on the price and required return. By being willing to trade more stock for a given market price, the \( G \) traders assist the non-\( G \) traders in absorbing the supply of the risky asset.
The corollary below follows immediately from the expression for \( a_j(G_j) \) in the above proposition:

**Corollary 1.**

1. \( a_j(G_j) \) decreases in \( G_j \), with \( a_j(0) = v_0 \); and
2. \( a_j(G_j) \) increases in \( \rho \).

Thus, the price of the \( j \)th basic security has a risk premium, and this premium decreases in \( G_j \), the direct utility of trading the \( j \)th basic security and in \( 1 - \rho \), the proportion of \( G \) traders. As the utility from trading grows without bound, the risk premium on the security decreases. Hence, \( G \) traders reduce risk premia and required returns on the security.

It is instructive to calculate the expected profits of the \( G \) traders. From Equations (4) and (5) and \( P1 \), the expected profits earned by a non-\( G \) and \( G \) trader, respectively, in the \( j \)th basic security, are given by:

\[
\begin{align*}
E\Pi_{NG,j} &= E[X_{NG,j}(P_j) \tilde{R}_j] = E\left[ \frac{\gamma a_j(G_j) (\xi_j + \xi_j)}{\gamma v_0} - (\theta_j + \gamma a_j(G_j) (\xi_j + \xi_j)) \right] \\
E\Pi_{G,j} &= E[X_{G,j}(P_j) \tilde{R}_j] = E\left[ \frac{\gamma a_j(G_j) (\xi_j + \xi_j)}{\gamma (v_0_G/G_j)^{1/2}} (\tilde{\theta}_j + \gamma a_j(G_j) (\xi_j + \xi_j)) \right] \\
&= \frac{\gamma a_j(G_j)^2}{v_0_G/G_j^{1/2}} \left( \xi_j + v_j \right).
\end{align*}
\]

Note that \( E\Pi_{G,j} \) > \( E\Pi_{NG,j} \). Thus, \( G \) traders earn greater expected profits than non-\( G \) traders. This simply emanates from the notion that, in effect, they trade more aggressively to capture the risk premium.

Of course, it follows that \( G \) traders bear more risk than non-\( G \) traders. Indeed, it is easy to show that the expected utility of wealth is smaller for \( G \) traders relative to non-\( G \) traders, that is \( E[-\exp(-\gamma W_t)] \) for the \( i \)th \( G \) trader is less than \( E[-\exp(-\gamma W_{t1})] \) for the \( i \)th non-\( G \) trader[13]. Our result is similar to that of Kyle and Wang (1997), who show that overconfidence acts as a commitment to trade aggressively in a strategic environment and thus results in greater expected profit for overconfident agents relative to that for rational agents. In our setting, agents who obtain direct utility from trading obtain a greater expected profit than neoclassical agents because the increased position size of the former class of agents increases their average gain, relative to that for the latter class. Like De Long et al. (1991), who obtain a similar result on the expected profits of overconfident vs rational traders, our result also challenges the notion that \( G \) traders would not exist because they would persistently lose money in financial markets.

### 2.2 Volatility

It follows from \( P1 \) that the price and return volatilities of the \( j \)th basic security are:

\[
\text{Var}(P_j) = (\gamma a_j(G_j))^2 v_j,
\]
Express the return of the market portfolio as $R_M = \sum_{j=1}^{N+K} (\xi_j + \tilde{z}_j) \tilde{R}_j$. We can compute the return volatility of the market portfolio. The computation of this variance, which is tedious, is provided in Appendix 2:

$$\text{Var}(\tilde{R}_M) = \sum_{j=1}^{N+K} \left[ \left( v_0 + 4\gamma^2 a_j(G_j)^2 v_{z_j} \right)^2 + v_0 v_{z_j} + 2\gamma^2 a_j(G_j)^2 v_{z_j}^2 \right].$$

(6)

Corollary 1 then implies the following results on the volatilities:

**Corollary 2.**

1. The individual security’s volatilities, $\text{Var}(P_j)$ or $\text{Var}(\tilde{R}_j)$, decrease in $G_j$ and increase in $\rho$; and
2. the aggregate volatility of the market portfolio, $\text{Var}(\tilde{R}_M)$, decreases in $G_j \forall j$ and increases in $\rho$.

These volatilities decrease in $1-\rho$, the mass of $G$ traders (and, of course, decrease in $G_j$). This is because $G$ traders behave like *de facto* liquidity providers[14]. When prices increase (drop), they sell (buy) more relative to the non-$G$ traders.

An empirical proxy of $\rho$ is the percentage holdings of institutional investors. The time trend in past decades is that institutional holdings have increased (e.g. Chordia et al., 2011). At the same time, there is evidence that firms’ individual volatilities have increased (Campbell et al., 2001). Our evidence accords with these stylized facts[15].

Note that the results in Corollary 2 related to $\rho$ hold only if $G_j > 0$. If $\forall j G_j = 0$, then $a_j(G_j) = v_0$ (see Corollary 1). The volatilities become $\text{Var}(P_j) = (\gamma v_0)^2 v_{z_j}$, $\text{Var}(\tilde{R}_j) = v_0 + (\gamma v_0)^2 v_{z_j}$ and:

$$\text{Var}(\tilde{R}_M) = \sum_{j=1}^{N+K} \left[ \left( v_0 + 4\gamma^2 v_0^2 v_{z_j} \right)^2 + v_0 v_{z_j} + 2\gamma^2 v_0^2 v_{z_j}^2 \right],$$

where the last equality follows from Equation (6). These volatilities are independent of $\rho$.

It is worth noting that in our setting, expected returns represent premia required to hold the positive mean supply $\xi_j$ in equilibrium. On the other hand, volatility is driven by fluctuations in premia required to hold the $z_j$ shocks to supply (which are random). The presence of $G$ traders mitigates both kinds of premia, so that these agents lower required returns and also lower volatility. In the next section, we formally derive results on covariance risk pricing.

### 2.3 The pricing of covariance risk

We now turn to how the $G$ traders affect $\beta$ pricing. We can compute the covariance between the returns of the $j$th basic security and the market portfolio (the computation, which is tedious, is provided in Appendix 2):

$$\text{Cov}(\tilde{R}_j, \tilde{R}_M) = \text{Cov} \left( \tilde{R}_j, \sum_{j=1}^{N+K} (\xi_j + \tilde{z}_j) \tilde{R}_j \right) = \text{Cov} \left( \tilde{R}_j, (\xi_j + \tilde{z}_j) \tilde{R}_j \right)$$

$$= v_0 \xi_j + 2\gamma^2 a_j(G_j)^2 v_{z_j} \xi_j.$$  

(7)
\[
E(\tilde{R}_j) = \gamma a_j(G_j) \tilde{\xi}_j = \gamma a_j(G_j) \text{Var}(\tilde{R}_M) \frac{1}{\nu_{\theta_j} + 2\gamma^2 a_j(G_j) v_{\xi_j}} \beta_{jM},
\]
(8)

where \( \beta_{jM} = (\text{Cov}(\tilde{R}_j, \tilde{R}_M))/(\text{Var}(\tilde{R}_M)) \).

Let:
\[
\lambda_j = \left( (\gamma a_j(G_j) \text{Var}(\tilde{R}_M) \right) / (\nu_{\theta_j} + 2\gamma^2 a_j(G_j)^2 v_{\xi_j})
\]
denote the slope of the relation between \( E(\tilde{R}_j) \) and \( \beta_{jM} \). The following proposition describes the comparative static for \( \lambda_j \) with respect to \( G_j \):

**P2**

1. Consider two basic securities, \( j \) and \( j' \), with \( \nu_{\theta_j} = \nu_{\theta_{j'}}, v_{\xi_j} = v_{\xi_{j'}}, \) but \( G_j > G_{j'} \). Then, \( \lambda_j < \lambda_{j'} \). The basic security with very large \( G_j \) (i.e. \( G_j / \gamma^2 > v_{\theta_j} \)) has \( \lambda_j \downarrow 0 \); and

2. \( \lambda_j \) increases in \( \rho \).

This proposition indicates that high \( G_j \) (or low \( \rho \) can lead to low \( \lambda_j \), and, therefore, attenuate the predictive power of \( \beta \)'s. Particularly, \( \lambda_j \downarrow 0 \) for stocks with very large \( G_j \) (i.e. \( G_j / \gamma^2 > v_{\theta_j} \)). In this extreme case, \( \beta \)'s lose power in explaining stock return completely. The intuition is that \( G \) traders, who obtain direct utility from “betting” on risky assets, reduce the equilibrium risk premia in these assets.

We now use \( P1 \) and Equation (7) to express the \( j \)th basic security's expected return as:

\[
E(\tilde{R}_j) = \gamma a_j(G_j) \tilde{\xi}_j = \gamma \text{Cov}(\tilde{R}_j, \tilde{R}_M) - \gamma \left[ \nu_{\theta_j} + 2\gamma^2 a_j(G_j)^2 v_{\xi_j} - a_j(G_j) \right] \tilde{\xi}_j
\]
(9)

where the last term, \(-\gamma \left[ \nu_{\theta_j} + 2\gamma^2 a_j(G_j)^2 v_{\xi_j} - a_j(G_j) \right] \tilde{\xi}_j\), is referred to as the \( \beta \)-adjusted expected return.

**P3**

1. The \( \beta \)-adjusted expected return of the \( j \)th basic security, \(-\gamma \left[ \nu_{\theta_j} + 2\gamma^2 a_j(G_j)^2 v_{\xi_j} - a_j(G_j) \right] \tilde{\xi}_j\), is negative.

2. Consider two basic securities, \( j \) and \( j' \), with \( \nu_{\theta_j} = \nu_{\theta_{j'}}, v_{\xi_j} = v_{\xi_{j'}}, \tilde{\xi}_j = \tilde{\xi}_{j'} \), but \( G_j > G_{j'} \). Then, the \( \beta \)-adjusted expected return of the \( j \)th basic security is lower than that of the \( j \)th basic security.

3. The \( \beta \)-adjusted expected return of the \( j \)th basic security increases in \( \rho \).

Thus, \( G \) traders cause securities to on average yield lower returns on a risk-adjusted basis. To see the intuition, note that since \( G \) traders get direct utility from trading, they are willing on average to pay more than rational investors for absorbing a given amount of supply, leading to lower returns than those naturally induced by risk premia.

### 2.4 Trading volume

We now examine trading volume within our model. We aim to ascertain how trading volume is influenced by the presence of agents who derive direct utility from trading, and to
investigate how volume might be associated with required returns on risky assets. It follows from Equation (4) and \( P1 \) that the \( i \)th non-\( G \) trader’s trade equals:

\[
X_{NG,i}(P_j) - (\xi_j + \tilde{z}_j) \sim N \left[ \frac{a_i(G_j)}{v_{\theta_j}} - 1 \right] \xi_j, \left( \frac{a_i(G_j)}{v_{\theta_j}} - 1 \right)^2 v_{\tilde{z}_j}. \tag{10}
\]

From Corollary 1, \( \left(\left(\frac{a_i(G_j)}/(v_{\theta_j})\right)-1\right)\xi_j < 0 \). Therefore, non-\( G \) traders on average take a short position in the \( j \)th basic security.

It follows from Equation (5) and \( P1 \) that the \( i \)th \( G \) trader’s trade equals:

\[
X_{G,i}(P_j) - (\xi_j + \tilde{z}_j) \sim N \left[ \frac{a_i(G_j)}{v_{\theta_j} - G_j/\gamma^2} - 1 \right] \xi_j, \left( \frac{a_i(G_j)}{v_{\theta_j} - G_j/\gamma^2} - 1 \right)^2 v_{\tilde{z}_j}. \tag{11}
\]

It is straightforward to show that \( \left(\left((a_i(G_j))/(v_{\theta_j} - G_j/\gamma^2)\right)-1\right)\xi_j > 0 \) Therefore, \( G \) traders on average take a long position in the \( j \)th security.

The expected trading volume is given by half the sums of the expected absolute changes in each type of agent’s position via trading in the market for the \( j \)th basic security. Using Equations (10) and (11), we can express the total expected trading volume in the basic security as:

\[
T_j \equiv 0.5 \rho E\left[|X_{NG,i}(P_j) - (\xi_j + \tilde{z}_j)|\right] + 0.5(1 - \rho) E\left[|X_{G,i}(P_j) - (\xi_j + \tilde{z}_j)|\right]. \tag{12}
\]

In Figure 1, we plot \( T_j \) as a function of \( \rho \) using the parameter values \( \gamma = 0.5, v_{\theta_j} = 1, \xi_j = 1 \) and \( v_{\tilde{z}_j} = 0.1 \). We use two different values for \( G_j \); the higher value is close to \( G_j \)’s upper limit.

**Notes:** This figure plots the expected trading volume \( T_j \) (see Equation (12)) as a function of \( \rho \), for two different values of \( G_j \). We assume the parameter values \( \gamma = 0.5, v_{\theta_j} = 1, \xi_j = 1 \) and \( v_{\tilde{z}_j} = 0.1 \).
As can be seen, volume is maximized for an interior value of $\rho$, because agent heterogeneity is needed for high volume. We also observe that the peak level of volume is much higher when $G_j$ is closer to its upper bound, as, close to that value, the trading activity of $G$ traders tends to increase sharply.

The following general result is readily derived from the expression for $T_j$:

Corollary 3. The expected trading volume, $T_j$, increases in $G_j$.

Corollary 3 indicates that stocks in which agents have a greater level of utility from trading exhibit greater trading volume. Intuitively, $G$ traders trade such stocks intensely, which, in turn, implies that non-$G$ traders take the opposite side of these positions in equilibrium, leading to high levels of volume. Also, since the $\beta$-adjusted expected return is more negative, the greater is $G_j$ ($P3$), our analysis indicates that, ceteris paribus, stocks with high volume (i.e. high $G_j$ stocks) will earn low average returns on a risk-adjusted basis. This is consistent with the negative relation between trading volume and required returns documented, for example, in Datar et al. (1998) and Brennan et al. (1998).

In a complementary explanation for volume premia, Baker and Stein (2004) argue that high volume implies high sentiment, and, under short-selling constraints, extreme optimism, that is subsequently reversed out. In contrast, our rationale for the link between expected returns and volume does not rely on short-selling constraints. Specifically, in the cross-section, high trading implies a high $G$, and consequently, a lower risk premium in the cross-section. Thus, our model does not predict intense buying in high volume stocks, whereas Baker and Stein’s (2004) model does. Based on Merton (1987), who argues that some (possibly, retail) investors might invest only in the most visible stocks, visibility (as measured by analyst following and brand visibility) might be a reasonable proxy for $G_j$. Our analysis suggests that such proxies will be associated with high volume and low average returns. In the next subsection, we consider another proxy for $G_j$, the volatility of the underlying asset’s cash flows.

### 2.5 $\beta$, idiosyncratic volatility and expected returns

We now model a situation where the utility from trading an asset depends on its payoff volatility. This assumption is motivated in part from work such as Linnet et al. (2012), Preuschoff et al. (2006) and Fiorillo et al. (2003), which indicates that the pleasure from gambling emanates from the volatility surrounding the final outcome and not from the expectation of profit. Further, Kumar (2009) empirically demonstrates that retail investors are attracted to volatile companies[16]. Intuitively, this captures the notion that agents prefer to gamble on companies with fluctuating payoffs like Alibaba, rather than on companies with steadier cash flows like Consolidated Edison (a utility company). Accordingly, we model the dependence of $G_j$ on firms’ cash flow volatilities as follows. Denote $\nu_\theta = \max(\nu_\theta)$ and $\nu_\delta = \min(\nu_\delta)$, and let:

$$G_j = G(\nu_\theta) = b \left( \nu_\theta - \nu_\delta \right),$$

where $b > 0$. In this specification, $G(\nu_\delta) = 0$ so the firm with the lowest cash flow volatility provides no direct utility from trading. We let $b$ take the highest level so the firm with the highest cash flow volatility $\nu_\theta$ still satisfies the requirement $G_j < \gamma^2 \nu_\theta$ from our earlier analysis; thus, $b = \gamma^2 (\bar{\nu}_\theta - \nu_\delta) - 1$, where $\gamma > 0$ is a sufficiently small positive constant. We finally assume that there is a sufficiently high mass of $G$ traders. Specifically:

$$\rho < \min \left( \frac{\nu_\delta}{\bar{\nu}_\theta}, 1 - \frac{\nu_\delta}{\nu_\theta - \nu_\delta} \right).$$
We then have the following lemma:

**Lemma 1.**

(1) The expected return of the \(j\)th basic security, \(E(\tilde{R}_j) = \gamma a_j(G_j) \bar{z}_j\) (see Equation (8)), decreases in \(v_{\theta_j}\); and

(2) *ceteris paribus*, the basic security with a higher \(v_{\theta_j}\) also has a higher \(\beta_j\).

This lemma suggests that if there is a significant mass of \(G\) traders (low \(\rho\)), then there can be a negative cross-sectional relation between \(\beta\) and expected return. The intuition is following. Risk aversion tends to cause positive pricing of covariance risk. But, if the mass of \(G\) traders is high, the strong volatility preference in the market more than offsets the effect of risk aversion, which causes \(\beta\) to be negatively priced. Our result is consistent with negative \(\beta\) pricing documented in some studies (Bali et al., 2016 (BBMT hereafter); Frazzini and Pedersen, 2014). Indeed, BBMT show that this phenomenon is more pronounced for stocks with smaller institutional holdings, consistent with our notion that \(G\) traders (likely non-institutions) facilitate negative \(\beta\) pricing.

We next show that under reasonable conditions, our analysis also accords with Ang et al. (2006), who demonstrate a negative cross-sectional relation between idiosyncratic volatility and average returns. For the \(j\)th basic security, if one runs a time series regression of \(\tilde{R}_j\) against \(\tilde{R}_M\), then the variance of the residual, which we refer to as the square idiosyncratic volatility (or simply IVOL), equals:

\[
IVOL_j = \frac{\text{Var}(\tilde{R}_j) - \frac{\text{Cov}(\tilde{R}_j, \tilde{R}_M)^2}{\text{Var}(\tilde{R}_M)}}{\text{Var}(\tilde{R}_M)}.
\]  

We let \(v_{\theta_j}\) vary while holding other exogenous parameters constant. Intuitively, the return IVOL should be positively related with the cash flow volatility measured by \(v_{\theta_j}\). It turns out in our model that this is true for “typical securities,” i.e., those with small \(\text{Var}(\tilde{R}_j)/\text{Var}(\tilde{R}_M)\). This implies that their value variation is relatively low compared to the value variation of the market portfolio (in a sense formalized in Appendix 1). The following lemma formalizes this observation:

**Lemma 2.** Consider two typical basic securities, \(j\) and \(j'\), with \(v_{z_j} = v_{z_{j'}}\), \(\bar{z}_j = \bar{z}_{j'}\) but \(v_{\theta_j} > v_{\theta_{j'}}\). Then, the idiosyncratic volatility \(IVOL_j > IVOL_{j'}\).

We then have the following proposition:

**P4.** For \(G_j = G(v_{\theta_j})\) given in Equation (13), the \(\beta\)-adjusted expected return,

\[-\gamma[v_{\theta_j} + 2\sigma^2 a_j(G_j)^2 v_{z_j} - a_j(G_j)]\bar{z}_j, \]

decreases in \(v_{\theta_j}\).

Lemma 2 and P4 imply that for typical basic securities, there is a negative relation (induced by \(v_{\theta_j}\)) between IVOL and the \(\beta\)-adjusted expected return. This is broadly consistent with Ang et al. (2006), who show that stocks with high 17 idiosyncratic volatility earn lower average returns. Proxying for total volatility by \(v_{\theta_j}\), our analysis also accords with Baker and Haugen (2012), who show that low risk stocks outperform high risk stocks in the vast majority of international equity markets.

The above analysis indicates that total volatility is negatively priced in the cross-section. However, in aggregate, risk is positively priced. To see this, note from P1 that the return of the market portfolio (over the risk-free interest rate which is normalized to be zero) is given by:

\[
\tilde{R}_M = \sum_{j=1}^{N+K} (\bar{z}_j + \tilde{z}_j) \tilde{R}_j = \sum_{j=1}^{N+K} [(\bar{z}_j + \tilde{z}_j) (\tilde{\theta}_j + \gamma a_j(G_j) (\bar{z}_j + \tilde{z}_j))].
\]
The following proposition can readily be derived:

$P5$. The market risk premium, $E(\hat{R}_M)$, is positive.

Thus, our model is consistent with the negative pricing of volatility in the cross-section, but a positive pricing of risk in the aggregate (Haugen and Baker, 2010; Ang et al., 2006; Mehra and Prescott, 1985).

2.6 Back to the original securities
The previous analysis focused on the basic securities for tractability. We now show that our main results carry over to the original securities. We can use Equation (2) to reconstruct the original securities. Specifically, consider the $j$th ($j = 1, \ldots, N$) non-factor original security as a portfolio of $V_j$ units of the risk-free asset, $\beta_{jk}$ units of the $N+k$th basic security (the $k$th factor, $\tilde{f}_k$) and one unit of the $j$th basic security (the $j$th residual, $\tilde{e}_j$). Denote the supply of the $j$th non-factor original security as $S_j$, and the supply of the $N+k$th original security as $S_{N+k}$. Note that the $j$th non-factor original security includes $\beta_{jk}$ units of the $k$th factor $\tilde{f}_k$ and one unit of $\tilde{e}_j$. Then, the supply of the $j$th ($j = 1, \ldots, N$) basic security is given

$S_j = x_j + \tilde{z}_j$, and the supply of the $N+k$th basic security (the $k$th factor) is given

$S_{N+k} = \sum_{j=1}^{N} S_j \beta_{jk} = \tilde{t}_{N+k} + \tilde{z}_{N+k}$.

Lemma 3. Trading the original securities is equivalent to trading the basic securities in that the price of the $j$th ($j = 1, \ldots, N$) non-factor original security, denoted by $P(\tilde{V}_j)$, is a linear combination of the prices of the basic securities:

$P(\tilde{V}_j) = \tilde{V}_j + \sum_{k=1}^{K} (\beta_{jk}P_{N+k}) + P_j$,

and the price of the $N+k$th ($k = 1, \ldots, K$) original security is given by $P(V_{N+k}) = P_{N+k}$, $P_j$ and $P_{N+k}$ are given in $P1$.

Now, note from $P1$ that the return of the $j$th non-factor original security can be expressed as:

$\tilde{R}_j = \sum_{k=1}^{K} \beta_{jk} [\tilde{a}_{N+k} + \gamma a_{N+k}(G_{N+k})(\tilde{t}_{N+k} + \tilde{z}_{N+k})] + [\tilde{\theta}_j + \gamma a_j(G_j) (\tilde{t}_j + \tilde{z}_j)]$. (17)

The expected return of the original security is given by:

$E(\tilde{R}_j) = \sum_{k=1}^{K} \beta_{jk} [\gamma a_{N+k}(G_{N+k})\tilde{t}_{N+k}] + [\gamma a_j(G_j) \tilde{t}_j]$. (18)

The expected return in our model takes a similar form as multi-factor models such as ICAPM (Merton, 1973) and APT (Ross, 1976). It follows from Corollary 1 that the expected return decreases in $G_j$. Let $\forall j, k \beta_{jk} \geq 0$. The expected return also decreases in $G_{N+k}$, and increases in $\rho$.

The volatility of the $j$th original security is given by:

$\text{Var}(\tilde{R}_j) = \sum_{k=1}^{K} \beta_{jk}^2 \left[ v_{N+k} + (\gamma v_{N+k})^2 v_{N+k} \right] + [v_{\theta_j} + (\gamma a_j(G_j))^2 v_{\theta_j}]$. 

It follows from Corollary 1 that the volatility decreases in $G_j$ and $G_{N+k}$ $\forall k = 1, \ldots, K$, and increases in $\rho$. 

232
Can \( \beta_{jM} \) predict returns? One can estimate \( \beta_{jM} \) by regressing \( \tilde{R}_j \) against the return on the market portfolio. Note that the market portfolio of the original securities is just a reshuffle of the basic securities, and is therefore identical to the market portfolio of the basic securities. It follows from Equations (7) and (17) that the market \( \beta \) of the \( j \)th original security is given by:

\[
\beta_{jM} = \frac{\sum_{k=1}^{K} \beta_{jk} \left[ v_{0N+k} + 2\gamma^2 a_{N+k}(G_{N+k})^2 v_{2N+k} \right] \xi_{N+k}}{\text{Var} \left( \tilde{R}_M \right)}.
\]

(19)

Write the expected return of the original security (see Equation (18)) as:

\[
E \left( \tilde{R}_j \right) = \lambda_j \beta_{jM},
\]

where:

\[
\lambda_j = \frac{\sum_{k=1}^{K} \beta_{jk} \left[ \gamma a_{N+k} (G_{N+k})^2 \xi_{N+k} \right] + \left[ \gamma a_j (G_j) \xi_j \right]}{\sum_{k=1}^{K} \beta_{jk} \left[ v_{0N+k} + 2\gamma^2 a_{N+k}(G_{N+k})^2 v_{2N+k} \right] \xi_{N+k} + \left[ v_{0} + 2\gamma^2 a_j (G_j)^2 v_{2} \right] \xi_j}} \text{Var} \left( \tilde{R}_M \right).
\]

We can use a similar analysis as in the proof of P2 to show that the original security with a higher \( G_j \) has a lower \( \lambda_j \) if \( \forall j, k, a_{N+k} (G_{N+k})^2 v_{2N+k} \geq a_j (G_j) v_2 \) (e.g. all basic securities have the same \( v_0 \) and \( v_2 \); but the factors have a lower \( G \) than the residuals, i.e. \( G_{N+k} \leq G_j \)). \( \lambda_j \) can be as low as zero.

Use Equations (18) and (19) to write the expected return of the \( j \)th original security as:

\[
E \left( \tilde{R}_j \right) = \gamma \text{Var} \left( \tilde{R}_M \right) \beta_{jM} + AR_j,
\]

where:

\[
AR_j = -\gamma \sum_{k=1}^{K} \beta_{jk} \left[ v_{0N+k} + 2\gamma^2 a_{N+k}(G_{N+k})^2 v_{2N+k} - a_{N+k}(G_{N+k}) \right] \xi_{N+k} - \left[ v_0 + 2\gamma^2 a_j (G_j)^2 v_2 - a_j (G_j) \right] \xi_j
\]

represents the \( \beta \)-adjusted expected return. We can use a similar analysis as in the proof of P3 to show that the \( \beta \)-adjusted expected return, \( AR_j \), is negative, and decreases in \( G_j \).

Note from Equation (2) that only the \( j \)th original security includes one unit of \( j \)th basic security. This implies that when an agent trades one unit of the \( j \)th original security, he trades one unit of \( j \)th basic security at the same time. Therefore, the trading volume of the \( j \)th original security is identical to that of the \( j \)th basic security given in Equation (12), which increases in \( G_j \) (see Corollary 3). Since the \( \beta \)-adjusted expected return, \( AR_j \), is more negative, the greater is \( G_j \) (see above analysis), ceteris paribus, stocks with high volume (i.e. high \( G_j \) stocks) will earn low average returns on a risk-adjusted basis.

Suppose \( G_j = G(v_0) \), given in Equation (13). We can use a similar analysis as in the proof of Lemma 1 to show that the expected return of the \( j \)th original security’s expected return, \( E(\tilde{R}_j) \) (see Equation (18)), decreases in \( v_0 \), while the original security with a higher \( v_0 \) also has a higher \( \beta_{jM} \) (see Equation (19)). Thus, there can be a negative cross-sectional relation (induced by \( v_0 \)) between \( \beta \) and the expected return.

One can estimate IVOL \( j \) by regressing \( \tilde{R}_j \) on the market portfolio’s returns. We can use a similar analysis as in the proof of Lemma 2 to show that the original security with a higher \( v_0 \)
also has a higher IVOL, and use a similar analysis as in the proof of P4 to show that the β-adjusted expected return, AR, decreases in ng. Thus, there can be a negative relation (induced by ng) between IVOL and the β-adjusted expected return.

Our result in P5 on the market risk premium being positive continues to hold for the original securities. The reason is that the market portfolio of the original securities is just a reshuffle of the basic securities, and is therefore identical to the market portfolio of the basic securities.

All of the preceding analysis indicates that results for the basic securities generally carry over to the original securities. Our work suggests the following untested empirical implications, which rely on the premise that stocks with more retail traders are likely to also have more G traders. We predict that ceteris paribus, stocks with proportionally more retail traders are less volatile and more actively traded. The cross-sectional (negative) relation between volatility and future returns is also likely to be more evident in stocks actively traded and held by retail investors.

3. Comparing to the economy with no or partial presence of G traders

We now compare the equilibria with complete absence of the G traders and presence of the G traders in some, but not all, securities. For simplicity and ease of exposition, the analysis in this section is focused on the basic securities, but the results carry over in a straightforward manner to the original securities (see Appendix 3). Further, we revert to the case of exogenous Gj in this section (and in the remainder of the paper), but the central points we wish to make do not depend on whether Gj is a function of volatility.

3.1 Comparing to the economy with no G traders

Consider two economies. In the first economy, all agents are non-G traders, while in the second, all are G traders. For a variable η in the basic economy, we use ηAG and ηANG to indicate its counterpart in the all-G and all-non-G economies, respectively.

The first economy, the all-non-G economy, is equivalent to the basic economy with ρ = 1 and ajGj/C0/C1 = ng. Using a derivation similar to that for P1, we can show that the price and return of the jth basic security are given by:

\[ P_{j,AN} = -\gamma v_0 \left( \xi_j + \tilde{z}_j \right), \] (20)

\[ \tilde{R}_{j,AN} = \tilde{\theta}_j + \gamma v_0 \left( \xi_j + \tilde{z}_j \right). \] (21)

The return on the market portfolio is:

\[ \tilde{R}_{M,AN} = \sum_{j=1}^{N+K} \left( \xi_j + \tilde{z}_j \right) \tilde{R}_{j,AN}. \]

Similar to Equation (7), the covariance between the returns of the basic security and the market portfolio is given by:

\[ \text{Cov} \left( \tilde{R}_{j,AN}, \tilde{R}_{M,AN} \right) = \text{Cov} \left( \tilde{R}_{j,AN}, \left( \xi_j + \tilde{z}_j \right) \tilde{R}_{j,AN} \right) = v_0 \xi_j + 2\gamma^2 v_0^2 v_0 \xi_j. \]

Then, the expected return on the basic security can be expressed as:

\[ E \left( \tilde{R}_{j,AN} \right) = \gamma v_0 \xi_j = \frac{\gamma v_0 \text{Var} \left( \tilde{R}_{M,AN} \right)}{v_0 + 2\gamma^2 v_0^2 v_0}, \]
where:

\[ \beta_{j\mathcal{M}}^{A,NG} = \frac{\text{Cov}(\tilde{R}_j^{A,NG}, \tilde{R}_M^{A,NG})}{\text{Var}(\tilde{R}_M^{A,NG})}. \]

Let:

\[ \lambda_j^{A,NG} = \left( \gamma v_0 \text{Var}(\tilde{R}_M^{A,NG}) / \left( v_0 + 2\gamma v_0 v_j \right) \right) \]

denote the slope of the relation between \( E(\tilde{R}_j^{A,NG}) \) and \( \beta_{j\mathcal{M}}^{A,NG} \). An immediate observation is that \( \lambda_j^{A,NG} > 0 \). Therefore, \( \beta \)'s still have power to predict stock returns. If \( v_j = 0 \), then \( \lambda_j^{A,NG} = \gamma \text{Var}(\tilde{R}_M^{A,NG}) \) is identical across all assets. In this case, \( \beta \)'s are the only predictive variable for expected returns.

The second economy, the all-\( G \) economy, is equivalent to the basic economy with \( \rho = 0 \) and \( a_j(G_j) = v_0 - G_j / \gamma^2 \). Using a derivation analogous to that for \( P1 \), we can show that the price and return of the \( j \)th basic security are given by:

\[ P_j^{A,G} = -\gamma (v_0 - G_j / \gamma^2) (\bar{z}_j + \tilde{z}_j), \]

(22)

\[ \tilde{R}_j^{A,G} = \tilde{\theta}_j + \gamma (v_0 - G_j / \gamma^2) (\bar{z}_j + \tilde{z}_j). \]

(23)

The return on the market portfolio is \( \tilde{R}_M^{A,G} = \sum_{j=1}^{N+K} (\bar{z}_j + \tilde{z}_j) \tilde{R}_j^{A,G} \).

Similar to Equation (7), the covariance between the returns of the basic security and the market portfolio is given by:

\[ \text{Cov}(\tilde{R}_j^{A,G}, \tilde{R}_M^{A,G}) = \text{Cov}(\tilde{R}_j^{A,G}, (\bar{z}_j + \tilde{z}_j) \tilde{R}_j^{A,G}) = v_0 \bar{z}_j + 2\gamma^2 (v_0 - G_j / \gamma^2)^2 v_j \bar{z}_j. \]

Then, the expected return on the basic security can be expressed as:

\[ E(\tilde{R}_j^{A,G}) = \gamma (v_0 - G_j / \gamma^2) \bar{z}_j = \frac{\gamma (v_0 - G_j / \gamma^2) \text{Var}(\tilde{R}_M^{A,G})}{v_0 + 2\gamma^2 (v_0 - G_j / \gamma^2)^2 v_j}, \beta_{j\mathcal{M}}^{A,G}. \]

where:

\[ \beta_{j\mathcal{M}}^{A,G} = \left( \text{Cov}(\tilde{R}_j^{A,G}, \tilde{R}_M^{A,G}) / \left( \text{Var}(\tilde{R}_M^{A,G}) \right) \right). \]

Let:

\[ \lambda_j^{A,G} = \left( \gamma (v_0 - G_j / \gamma^2) \text{Var}(\tilde{R}_M^{A,G}) / \left( v_0 + 2\gamma^2 (v_0 - G_j / \gamma^2)^2 v_0 \right) \right), \]

denote the slope of the relation between \( E(\tilde{R}_j^{A,G}) \) and \( \beta_{j\mathcal{M}}^{A,G} \).
We compare the two economies in the following proposition:

\[ P6. \] First, \( E(\tilde{R}^{A,NG}_{j}) > E(\tilde{R}^{A,G}_{j}) \) and \( \text{Var}(\tilde{R}^{A,NG}_{j}) > \text{Var}(\tilde{R}^{A,G}_{j}) \). Thus, the \( j \)th basic security has higher expected return and volatility in the all-non-\( G \) economy than in the all-\( G \) economy. Second, \( E(\tilde{R}^{A,NG}_{M}) > E(\tilde{R}^{A,G}_{M}) \) and \( \text{Var}(\tilde{R}^{A,NG}_{M}) > \text{Var}(\tilde{R}^{A,G}_{M}) \). Thus, the market portfolio has higher expected return and volatility in the all-non-\( G \) economy than in the all-\( G \) economy. Third, \( \lambda_{j}^{A,NG} > \lambda_{j}^{A,G} \). Thus, \( \beta \)s have more predictive power in the all-non-\( G \) economy than in the all-\( G \) economy.

In general, within the all-\( G \) economy, risk premiums (and volatilities) are attenuated because of the \( G \) traders’ penchant to buy low and sell high, which, in turn, attenuates the pricing of risk. This suggests the testable implication that economies which are dominated by retail investors should exhibit lower volatility and less evidence of covariance risk pricing. Indeed, under the assumption that retail investors are more likely to participate during periods of positive sentiment (Grinblatt and Keloharju, 2001; Lamont and Thaler, 2003), and less likely to participate during periods of negative sentiment, our analysis accords with Antoniou et al. (2016), who show that covariance risk pricing is more prevalent during periods of low sentiment and vice versa.

3.2 Comparing to the economy with partial presence of \( G \) traders

Consider a hybrid case in which \( G \) traders are present in the trading of the basic securities mimicking the \( K \) factors and the first \( N_{1} \) residuals. They are not present in the remaining \( N-N_{1} \) basic securities mimicking the remaining \( N-N_{1} \) residuals. Like in the basic economy, there are masses \( \rho \) and \( 1-\rho \) of non-\( G \) and \( G \) traders, respectively.

We continue to write the return of the \( j \)th basic security as \( \tilde{R}_{j} = \tilde{\theta}_{j} - \tilde{P}_{j} \). (For convenience, we use the same notation for prices in all securities, i.e. \( \tilde{P}_{j} \) for the \( j \)th security, even though prices of securities without \( G \) traders can take a different form from that in Section 2.1.) The market portfolio has a return \( \tilde{R}_{M} = \sum_{j=1}^{N+k} (\tilde{z}_{j} + \tilde{z}_{j}) \tilde{R}_{j} \).

For the \( j \)th basic securities mimicking the \( K \) factors and the first \( N_{1} \) residuals, our analysis in Section 2.3 still holds. Particularly, the expected return of the \( j \)th basic security takes the form:

\[ E\left(\tilde{R}_{j}\right) = \lambda_{j}\beta_{j,M}, \]

where \( \lambda_{j} = (\lambda a_{j}(G) \text{Var}(\tilde{R}_{M}))/(v_{0} + 2\gamma^{2}a_{j}(G)^{2}v_{z}) \) denotes the slope of the relation between \( E(\tilde{R}_{j}) \) and \( \beta_{j,M} \).

For the remaining \( N-N_{1} \) basic securities, there is only a mass \( \rho \) of non-\( G \) traders to clear the market. Using a similar analysis as in Section 2.1, we can show that for these securities:

\[ P_{j}^{\rho} = -\gamma(v_{0}/\rho)(\tilde{z}_{j} + \tilde{z}_{j}), \]

\[ \tilde{R}_{j}^{\rho} = \tilde{\theta}_{j} + \gamma(v_{0}/\rho)(\tilde{z}_{j} + \tilde{z}_{j}). \]

Here, we use the superscript \( \rho \) to indicate the \( N-N_{1} \) securities. Similar to Equation (7), the covariance between the returns of the basic security and the market portfolio is given by:

\[ \text{Cov}\left(\tilde{R}_{j}^{\rho}, \tilde{R}_{M}\right) = v_{0}\tilde{z}_{j} + 2\gamma^{2}(v_{0}/\rho)^{2}v_{z}\tilde{z}_{j}. \]
Then, the expected return of the basic security can be expressed as:

\[ E(\tilde{R}_j^o) = \gamma(v_0/\rho) \bar{\xi}_j = \frac{\gamma(v_0/\rho) \text{Var}(\tilde{R}_M)}{v_0 + 2\gamma^2(v_0/\rho)\bar{v}_j} \beta_{jM}^o = \lambda_j^o \beta_{jM}^o, \]

where:

\[ \beta_{jM}^o = \left( \text{Cov}(\tilde{R}_j, \tilde{R}_M) \right) / \left( \text{Var}(\tilde{R}_M) \right), \]

and:

\[ \lambda_j^o = \left( \gamma(v_0/\rho) \text{Var}(\tilde{R}_M) \right) / \left( v_0 + 2\gamma^2(v_0/\rho)^2\bar{v}_j \right), \]

denote the slope of the relation between \( E(\tilde{R}_j^o) \) and \( \beta_{jM}^o \). We then have the following proposition.

**P7.** Consider two basic securities, \( j \) and \( j' \), with \( v_0 = v_{0j}, v_z = v_{zj}, \bar{\xi}_j = \bar{\xi}_j' \), but the \( j \)th (\( j \)th) security is (is not) traded by \( G \) traders. Then, \( \lambda_j < \lambda_j' \). Thus, the presence of \( G \) traders reduces the predictive power of \( \beta \).

The above proposition implies that \( \beta \) pricing will be less evident in securities that are traded relatively more by \( G \) traders. Again, the notion is simply that \( G \) traders, via their more aggressive trading in securities where they are present, attenuate the pricing of risk. The above proposition indicates cross-sectional variation in risk pricing according to whether \( G \) traders (likely retail investors) are prevalent in a particular stock or not.

### 4. Stock market vs other gambling avenues

Agents in the real world have a large number of opportunities to indulge in risky gambles; for example, gambling in casinos or participating in lotteries. An interesting issue is the incentives to substitute across various ways to gamble in risky opportunities. Accordingly, in this section, we introduce a class of agents, a mass \( \rho^A \) of “\( G^A \) traders.” These agents may choose to either invest in the stock market, or gamble through an alternative venue, such as a casino or a lottery (but not both). As our focus in this section and the next is not on the cross-section, we focus on the case of a single risky security as representative of the stock market[17]. The payoff on this security follows a normal distribution with \( V \sim N(\overline{V}, \upsilon) \). Each \( G^A \) trader is endowed with \( M_i \) units of the risk-free asset. If a \( G^A \) trader chooses to gamble through the alternative venue, then his utility function takes the following form:

\[ U_{G^A}(W_{i1}, Y) = -\exp(-\gamma W_{i1} - 0.5Y^2G^A), \]

where \( W_{i1} = M_i + Y\tilde{\mu} \). \( \tilde{\mu} \) represents the payoff of this gambling opportunity (we normalize its mean to be zero). \( Y > 0 \) is a positive constant representing his fixed position. \( G^A \) represents the parameter that governs his utility derived from this gambling opportunity.
If a $G^4$ trader invests in the stock market, then he/she has the same utility function as that of the $G$ traders (see Equation (1)). We have that:

$$U_{G^4}(W_{i1}, X_i) = -\exp(-\gamma W_{i1} - 0.5G X_i^2),$$

where $W_{i1} = M_i + X_i (V - P)$, $X_i$ is his/her demand for the risky security and $P$ is the price of the risky security. It follows, after taking expectations over $V$, that:

$$E[U_{G^4}(W_{i1}, X_i)] = -\exp \left[ -\gamma M_i - \gamma X_i (V - P) + 0.5\gamma^2 (v - G/\gamma^2) X_i^2 \right]. \quad (24)$$

Let $\rho_1 \in [0, \rho^4]$ be the mass of $G^4$ traders who choose to invest in the stock market. It follows from an analysis similar to that in Section 2.1 that each non-$G$, $G$ and $G^4$ trader’s demands for the $j$th basic security are, respectively, given by:

$$X_{NG}(P) = \frac{V - P}{\gamma v}, \quad (25)$$

$$X_G(P) = \frac{V - P}{\gamma (v - G/\gamma^2)}, \quad (26)$$

$$X_{G^4}(P) = \frac{V - P}{\gamma (v - G/\gamma^2)}. \quad (27)$$

The per capita supply of the security equals $\bar{z} + \bar{x}$, which, as in the previous analysis, equals the endowment of each $G$ and non-$G$ trader. The market clearing condition, $\rho X_{NG}(P) + (1 - \rho) X_G(P) + \rho_1 X_{G^4}(P) = \bar{z} + \bar{x}$, then implies that the price of the $j$th basic security is given by:

$$P = \frac{V - \gamma a(G, \rho_1) (\bar{z} + \bar{x})}{\gamma}, \quad (28)$$

where $a(G, \rho_1) = 1/(\rho/v) + (1 - \rho + \rho_1)/(v - G/\gamma^2))$. Substituting for the $G^4$ trader’s optimal demand, $X_{G^4}(P)$ from Equation (27) (where $P$ is given in Equation (28)) into Equation (24), we obtain:

$$E[U_{G^4}(W_{i1}, X_i)] = -\exp \left[ -\gamma M_i - 0.5\gamma^2 (v - G/\gamma^2) X_i^2 \right]$$

$$= -\exp \left[ -\gamma M_i - 0.5\gamma^2 a(G, \rho_1) (\bar{z} + \bar{x}) \right]. \quad (29)$$

If the $G^4$ trader gambles in the alternative opportunity, it follows after taking expectations over $\bar{m}$ that:

$$E[U_{G^4}(W_{i1}, Y)] = -\exp \left[ -\gamma M_i + 0.5\gamma^2 Y^2 \right]. \quad (30)$$

The $G^4$ trader either invests in the stock market or gambles through the alternative venue based on a comparison between the ex ante expected utility from trading in the stock market (taking expectations over $\bar{z}$ in Equation (29)) and the expected utility from the alternative venue in Equation (30). One then obtains the following proposition:

$P8.$ In equilibrium, the mass $\rho_1$ of $G^4$ traders who trade in the stock market decreases in $G^4$. 

If the alternative gambling opportunity becomes more attractive (higher $G^A$), then less $G^A$ traders will trade in the stock market.

Using Equations (25)–(27), we can express the total expected trading volume in the basic security as:

$$T(\rho_1) \equiv 0.5\rho E\left[|X_{NG}(P)-(\xi+\tilde{z})|\right] + 0.5(1-\rho)E\left[|X_G(P)-(\xi+\tilde{z})|\right] + 0.5\rho_1E\left[|X_{G^A}(P)|\right]. \quad (31)$$

We then have the following result:

**Corollary 4.** The expected trading volume in equilibrium, $T(\rho_1)$, increases in $\rho_1$ and decreases in $G^A$.

We thus find that substituting across various avenues can affect trading volume in equities. Specifically, increased attractiveness of the alternative gambling opportunity causes a decrease in equilibrium equity volume. This result is consistent with Gao and Lin (2015). They show that equity trading volume in Taiwan decreases as the total jackpot of a major statewide lottery increases.

5. **A dynamic extension**

We now consider an intertemporal version of our setting where the utility from trading depends on past profits. Specifically, we model the notion that if an agent earns positive profits, he/she may derive greater utility from gambling in the stock market. This may happen because of the “house money effect” (Thaler and Johnson, 1990) or a direct physiological response to winning (Coventry and Constable, 1999). We show that the “excitement” created by positive profits can lead to an overreaction to mildly positive information and thus cause a “bubble” in stock prices. Our motivation is consistent with experimental arguments that emotional excitement can cause bubbles (Bellotti et al., 2010; Andrade et al., 2016); in our model, the pathway is that positive profits increase the “excitement” or utility from additional trading and thus cause an overreaction to mildly positive information. We also show that our analysis accords with the “leverage effect” in equities, (Black, 1976; Christie, 1982), wherein down markets are followed by increased volatility.

We assume that a risky security is traded at Dates $t = 0, 1, 2$ and 3. While we analyze a single security, without loss of generality, it can be viewed as any basic security defined in Section 2.1. For convenience, we assume that the supply of the risky security, $\xi > 0$, is drawn from the support $(0, 1)$ according to a cumulative density function with a variance $\text{Var}(\xi)$. Traders observe $\xi[18]$. At Date 3, the security pays off a liquidation dividend:

$$V = V + \tilde{\theta}_1 + \tilde{\theta}_2 + \tilde{\theta}_3,$$

where $V$ is a positive constant, which represents the expected dividend. The variables $\tilde{\theta}_t$, $t = 1, 2$ and 3 represent exogenous cash flow shocks, which are mutually independent and multivariate normally distributed with mean zero. $\tilde{\theta}_t$’s are public signals released at Dates $t = 1, 2$. Its prices are $P_t$ at Dates $t = 0, 1$ and 2.

There is a mass unity of identical agents who trade the risky security. At Date 0, they all hold an identical long position, in aggregate $\xi$, to clear the market. At Date 1, if $P_1 > P_0$, they make money at Date 1; otherwise, they do not. If $P_1 > P_0$, so that they make money at Date 1, then after Date 1 and before Date 2, with probability $\rho$, an agent becomes a non-$G$ trader; and with probability $1-\rho$, he becomes a $G$ trader. If $P_1 < P_0$ and the agents do not make money at Date 1, then all agents remain non-$G$ traders with a probability of unity.

The $i$th non-$G$ trader’s utility function is the standard exponential:

$$U(W_{i3}) = -\exp(-\gamma W_{i3}).$$
where $W_3$ is his final wealth, and $\gamma$ is a positive constant representing the absolute risk aversion coefficient. The $i$th $G$ trader’s utility function takes the form:

$$U_G(W_{i3}, C_{i3}) = -\exp(-\gamma W_{i3} - C_{i3})$$

$C_{i3}$ captures the extra (direct) utility from the act of trading. We let $C_{i3} = 0.5 G X_{i2}^2$, where $X_{i2}$ is the quantity of risky security he has after trading at Date 2, and $G$ is a positive constant. Thus, his utility function can be expressed as:

$$U_G(W_{i3}, X_{i2}) = -\exp(-\gamma W_{i3} - 0.5 G X_{i2}^2)$$

For convenience, we let $\tilde{y}_t's, t = 1, 2, and 3,$ have the same variance $\nu_\theta$. Let the price and return of the risk-free asset be 1. We then have the following result:

**P9.** There is an equilibrium characterized by the following prices:

- $P_0$ is given by:

$$P_0 = V + H_\theta - 2\gamma v_\theta \xi.$$  

The variable $H_\theta$ is uniquely determined by:

$$0 = \int_{H_\theta}^{\infty} \frac{\tilde{\theta}_1 - H_\theta + \gamma(v_\theta - a(G))\xi}{\exp(\gamma \xi \tilde{\theta}_1)} \left[ \rho \exp\left(-0.5\frac{(\gamma a(G)\xi)^2}{v_\theta}\right) \right] d\Phi\left(\frac{\tilde{\theta}_1}{\sqrt{v_\theta}}\right)$$

$$+ (1-\rho) \exp\left(-0.5\frac{(\gamma a(G)\xi)^2}{v_\theta - G/\gamma^2}\right) d\Phi\left(\frac{\tilde{\theta}_1}{\sqrt{v_\theta}}\right)$$

$$+ \int_{-\infty}^{H_\theta} \frac{\tilde{\theta}_1 - H_\theta}{\exp(\gamma \xi (\tilde{\theta}_1 - (v_\theta - a(G))\xi))} \exp\left(-0.5\frac{(\gamma v_\theta \xi)^2}{v_\theta}\right) d\Phi\left(\frac{\tilde{\theta}_1}{\sqrt{v_\theta}}\right),$$

where $a(G) = [(\rho/\nu_\theta) + (1-\rho)/(\nu_\theta - G/\gamma^2)]^{-1}$, and $\Phi(.)$ is the cumulative density function of standard normal distribution:

- If $\tilde{\theta}_1 > H_\theta$, then:

$$P_1 = V + \tilde{\theta}_1 - \gamma v_\theta \xi - \gamma a(G)\xi, \quad \text{and} \quad P_2 = V + \tilde{\theta}_1 + \tilde{\theta}_2 - \gamma a(G)\xi.$$  

- If $\tilde{\theta}_1 \leq H_\theta$, then:

$$P_1 = V + \tilde{\theta}_1 - 2\gamma v_\theta \xi, \quad \text{and} \quad P_2 = V + \tilde{\theta}_1 + \tilde{\theta}_2 - \gamma v_\theta \xi.$$  

Here is a sketch of the proof of this proposition (the formal proof is in Appendix 1). We use backward induction. There are three steps. In the first step, we study the equilibrium demands and prices at Dates 1 and 2 conditional on the event that at Date 1, $\tilde{\theta}_1 > H_\theta$ so that $P_1 > P_0$ (call this Regime 1). Note that in this regime, at Date 2, there is a mass $\rho (1-\rho)$ of non-$G$ traders ($G$ traders). At Date 1, an agent knows that he will be a non-$G$ trader ($G$ trader) with probability $\rho (1-\rho)$. In the second step, we study the equilibrium demands and prices at dates 1 and 2 conditional on the event that at Date 1, $\tilde{\theta}_1 \leq H_\theta$ so $P_1 \leq P_0$ (call this Regime 2). This step is simpler than the first step because all traders are non-$G$ traders. In the third step, we focus on Date 0, and derive the expressions for $P_0$ and the threshold $H_\theta$. 


There are two interesting results. The first relates to the price reaction for \( \tilde{\theta}_1 \) around the threshold \( H_\theta \):\[ P_1(\tilde{\theta}_1 \leq H_\theta) = V + \tilde{\theta}_1 - \gamma v_\theta \xi - \gamma a(G)\xi, \]
\[ P_1(\tilde{\theta}_1 > H_\theta) = V + \tilde{\theta}_1 - 2\gamma v_\theta \xi. \]

It is easy to show that:
\[ P_1(\tilde{\theta}_1 \leq H_\theta) - P_1(\tilde{\theta}_1 > H_\theta) = \gamma v_\theta \xi - \gamma a(G)\xi > 0, \]

because \( a(G) < a(0) = \nu_\theta \). This suggests a small \( \tilde{\theta}_1 \) (e.g. earnings) can induce a significant price movement. Another interpretation of this observation is that a relatively minor piece of news can cause substantial moves in prices.

The second result relates to long-run performance. If \( \tilde{\theta}_1 > H_\theta \), then the subsequent returns are:
\[ P_2 - P_1 = \tilde{\theta}_2 + \gamma v_\theta \xi, \quad \text{and} \quad V - P_2 = \tilde{\theta}_3 + \gamma a(G)\xi. \]

If \( \tilde{\theta}_1 \leq H_\theta \), then the subsequent returns are:
\[ P_2 - P_1 = \tilde{\theta}_2 + \gamma v_\theta \xi, \quad \text{and} \quad V - P_2 = \tilde{\theta}_3 + \gamma v_\theta \xi. \]

A comparison between these two cases suggests that if \( \tilde{\theta}_1 > H_\theta \), then there is long-run underperformance (lower \( V - P_2 \)) because \( a(G) < a(0) = \nu_\theta \). Thus, a minor piece of good news can cause securities to become dramatically overpriced and subsequently exhibit subpar returns.

Figure 2 plots the price paths conditional on the public announcement \( \tilde{\theta}_1 \). We assume the parameter values \( V = 5 \) and \( \nu_\theta = 1 \), \( \xi = 1 \), \( \gamma = 0.5 \), \( \rho = 0.5 \) and \( G = 0.2 \). The realizations of \( \tilde{\theta}_2 \) and \( \tilde{\theta}_3 \) are assumed to be zero, i.e. their mean. This implies that the threshold \( H_\theta = -0.388 \). Moving from the bottom to the top, each path in the figure represents a realization of \( \tilde{\theta}_1 \) from \(-1\) to \(0\) (step size = 0.025). \( \tilde{\theta}_1 \leq H_\theta \) for the paths indicated by \( \Delta s \), \( \tilde{\theta}_1 > H_\theta \) for the paths indicated by \( \ast s \).

We see that if \( \tilde{\theta}_1 \) is below the threshold \( H_\theta = -0.388 \), the price reaction to \( \tilde{\theta}_1 \) is non-positive. Once \( \tilde{\theta}_1 \) has surpassed the threshold \( H_\theta = -0.388 \), the price reaction becomes positive.

Particularly, consider the two paths bordering the hollow area. The south path is for \( \tilde{\theta} = -0.4 \). The north path is for \( \tilde{\theta} = -0.375 \). Although \( \tilde{\theta}_1 \) differs by only 0.025 across the two path groups, the price reactions are very different. On both paths, \( P_0 = 3.612 \). However, on the south path, \( P_1 = 3.6 \) so \( P_1 - P_0 = -0.012 \); on the north path, \( P_1 = 3.9583 \) so \( P_1 - P_0 = 0.3463 \). The difference in the price reaction, \( P_1 - P_0 \), equals 0.3583, which is more than 14 times the difference in \( \tilde{\theta}_1 \) (0.025).

The immediate return subsequent to the release of \( \tilde{\theta}_1 \), \( P_2 - P_1 = 0.5 \), is identical across all \( \tilde{\theta}_1 \) paths. But the long-run performance for the paths with \( \tilde{\theta}_1 > H_\theta \) indicated by \( \ast s \), \( V - P_2 = 0.1667 \), is lower than that for the paths with \( \tilde{\theta}_1 \leq H_\theta \) indicated by \( \Delta s \), \( V - P_2 = 0.5 \). This indicates long-run underperformance following a good public announcement. The underperformance is characteristic of bubble-like episodes in the stock market (such as the technology bubble of the 1990s, namely, Brunnermeier and Nagel, 2004), whereas the positive event (that creates the bubble) could be something as simple as good initial sales or earnings figures for the relevant sector.

More generally, the preceding analysis suggests a testable implication. Specifically, for stocks that are popular amongst retail investors, we predict a nonlinear response to positive news, that is, a small reaction to retail market announcements, but a disproportionately larger reaction to major (positive) announcements. Following the large positive announcements, these stocks should exhibit long-run reversals (conditional on the news).
We now turn to volatility and volume. Note that from an econometrician’s perspective, $\xi$ has a variance $\text{Var} (\xi)$. Further, trade in this model occurs at Date 2, since at Date 1, traders have identical and endowments. Let $T_2$ represent the expected volume at this date.

We then have the following results:

**Corollary 5.**

First, $\text{Var}(P_2 - P_1 | P_1 > P_0) = \text{Var}(P_2 - P_1 | P_1 \leq P_0)$ and $\text{Var}(V - P_2 | P_1 > P_0) < \text{Var}(V - P_2 | P_1 \leq P_0)$, so, overall, the volatility of price changes is high (low) when past stock performance (i.e. $P_1 - P_0$) is negative (positive). Second, $T_2(P_1 > P_0) > T_2(P_1 \leq P_0)$, so expected trading volume is high (low) when past stock performance (i.e. $P_1 - P_0$) is positive (negative).

Thus, our analysis accords with the leverage effect (Black, 1976; Christie, 1982; Aït-Sahalia et al., 2013), wherein down markets are followed by higher volatility and vice versa[19]. The analysis is also consistent with early work by Comiskey et al. (1987), as well as Chordia et al. (2007) (see also Karpoff, 1987), who show that volume is higher for stocks with high returns in the recent past. Intuitively, a drop in returns is accompanied by less trading interest from G traders, thus reducing liquidity provision[20], decreasing volume and increasing volatility.

**Notes:** This figure plots the price paths conditional on the public announcement $\theta_1$. We assume the parameter values $\bar{V} = 5$, $v_0 = 1$, $\bar{\xi} = 1$, $\gamma = 0.5$, $\rho = 0.5$ and $G = 0.2$. This implies that the threshold for the probabilistic conversion to a $G$ trader is $H_0 = -0.388$. The realizations of $\theta_2$ and $\theta_3$ are assumed to be zero, i.e., their mean. From the bottom to the top, each path represents a realization of $\theta_1$ from -1 to 0 (step size = 0.025). $\theta_1 \leq H_0$ for the paths indicated by $\Delta$s. $\theta_1 > H_0$ for the paths indicated by *s.

**Figure 2.** Price reaction and long-run performance
6. Conclusion

We present a model where agents derive direct utility from trading. Such traders attenuate $\beta$ pricing and volatility, while raising trading volume. The analysis thus accords with the lack of evidence consistent with covariance risk pricing (Fama and French, 1992), and a negative relation between average returns and volume (Datar et al., 1998). To the extent that agents who trade for entertainment tend to be retail investors, our setting also explains the rise of volatility as well as institutional (and a decrease in individual) holdings (Campbell et al., 2001; Malkiel and Xu, 1999). We next consider the neuro-psychological finding (Linnet et al., 2012; Fiorillo et al., 2003) that it is the ex ante volatility of the gambling outcome that is associated with the secretion of the pleasure-enhancing hormone dopamine. We thus assume that agents derive greater utility from trading more volatile stocks. The analysis then accords with negative volatility pricing in the cross-section (namely, Ang et al., 2006; Baker and Haugen, 2012), though volatility is still priced in the aggregate. Further, in a dynamic setting, when agents' utility from trading depends positively on past profit outcomes (Thaler and Johnson, 1990; Coventry and Constable, 1999), the equilibrium yields pricing “bubbles,” the leverage effect (Black, 1976; Christie, 1982) and greater volume in up relative to down markets (Karpoff, 1987).

Untested implications of our analysis are that cross-sectionally, stocks in economies where retail investors are dominant should exhibit weaker evidence of covariance risk pricing and greater share turnover. Such stocks should also exhibit disproportionate price reactions to moderately positive news announcements. Our work, however, raises many issues. First, it would be interesting to examine a fully dynamic model with exits and entry by agents who derive direct utility from trading. Second, it may be interesting to combine trading for entertainment and other investor biases, such as representativeness and overconfidence, and to examine the market equilibrium that results. Finally, the specific factors that influence how much utility is derived per unit trade (such as age, personality attributes) need to be considered in more depth within a theoretical setting. These and other issues are left for future research.

Notes


3. A recent article in the Wall Street Journal (“Let’s Be Honest: Are You an Investor or a Speculator?” by Jason Zweig, December 9, 2016) quotes “Barry Metzger, a senior vice president for trading services at Charles Schwab Corp.,” as saying that some active traders “believe that trading is fun and a part of who they are,” indicating that they view trading as a consumption good.


5. An issue is whether the mass of traders who trade for entertainment is sufficiently large to affect prices. On this issue, Barber et al. (2009) show that retail traders (who are more likely to trade for enjoyment than professionals), do have an impact on financial market prices.

6. Our paper is complementary to Friedman and Heinle (2016) and Luo and Subrahmanyam (2017), who consider a setting where agents derive direct utility or disutility from owning certain types of stocks. The utility there emanates from the signed position; for example, an environmentally conscious agent derives disutility from owning stocks (and utility from shorting stocks) in firms that heavily use coal. In contrast, motivated by the literature on gambling, direct utility in our model emanates from the unsigned quantity of trade.
7. Barrot et al. (2016) and Kaniel et al. (2008) demonstrate that individual investors do act as liquidity providers in financial markets.

8. Kumar (2009), Doran et al. (2012), Shefrin and Statman (2000) and Brunnermeier et al. (2007) also discuss skewness, but since our model has normally distributed payoffs it unfortunately cannot speak to skewness preference.

9. In a complementary and important view, Odean (1998) and Statman et al. (2006) show that overconfidence creates excessive trading volume. A distinguishing feature of our approach from theirs is that overconfidence increases price volatility whereas our agents who derive trading utility reduce volatility. In Daniel et al. (2001), overconfidence attenuates β pricing, just like in our paper. However, overconfidence is unlikely to explain phenomena such as negative pricing of risk, since there is no reason to believe that overconfident agents should be more attracted to risky assets.

10. In an alternative setting, this direct utility might occur on a per trade basis, and might depend on the frequency of trading. We abstract from such issues in our paper.

11. See Van Nieuwerburgh and Veldkamp (2010), Banerjee (2011), Daniel et al. (2001) and Caskey (2017) for other papers that use this approach. In Section 2.6, we show that our results also apply to the original securities but the exposition is more complicated.

12. As in earlier literature on exponential-normal models (e.g. Hong and Stein, 1999), price changes are viewed as synonymous with returns in this paper.

13. This result is obvious from the fact that G traders do not maximize \( E[-\exp(-\gamma W_{1})] \).

14. Kaniel et al. (2008) provide evidence that individual investors provide liquidity to institutions.

15. Note that our implication contrasts with overconfidence-based models; thus, Daniel et al. (1998) show that overconfidence increases volatility.

16. The analysis of Kumar (2009), in turn, is derived from the notion that unsophisticated agents are more attracted to lotteries (Rubenstein et al., 2002), and lotteries demonstrate extremely high variance and high skewness. As our model assumes normal distributions, it cannot speak to skewness, of course, and considers volatility instead.

17. Most of the results in the remainder of the paper readily generalize to multiple securities; details are available from the authors on request.

18. Restricting ourselves to the case of a single security and constraining the support to be positive simplifies the analysis. We have verified that generalizing to cases where agents trade multiple securities and face uncertain supply with unbounded support is possible, but we lose closed-form solutions and the presentation of the analysis becomes more complicated, without the promise of additional intuition.

19. While the term “leverage effect” refers to the notion that down markets imply increased leverage in market value terms and thus increased volatility, Figlewski and Wang (2000) argue that the asymmetric volatility–return relation is not linked to firm leverage and is a direct consequence of down markets, as in our analysis.

20. Indeed, Hameed et al. (2010) show measures of liquidity provision are higher in up markets relative to down markets.

21. If \( y \sim N(\gamma, \nu) \), then \( E[y] = \sqrt{2\phi((\gamma)/\sqrt{\nu}) + ((\gamma)/\sqrt{\nu})(1-2\Phi((-\gamma)/\sqrt{\nu}))} \), where \( \phi(\cdot) \) and \( \Phi(\cdot) \) denote the probability density function (pdf) and cumulative density function (cdf), respectively, of standard normal distribution.

References


## Appendix 1

Proof of Corollary 1: From $P1$, $a_j(G_j) = 1/((\rho)/(v_{\theta_j})+(1-\rho)/(v_0-G_j/\gamma^2))$. It is easy to show after taking derivatives that $a_j(G_j)$ decreases in $G_j$ and increases in $\rho$. Finally, $a_j(0) = 1/((\rho)/(v_{\theta_j})+(1-\rho)/(v_0)) = v_{\theta_j}$.

Proof of $P2$:

1. From Corollary 1, $a_j(G_j) < a_j(G_j)$ because $G_j > G_j$. Note that $\lambda_j = (\gamma a_j(G_j)\text{Var}(\tilde{R}_M))/\left((v_0 + 2\gamma^2 a_j(G_j)^2 v_z)\right)$. For $\lambda_j < \lambda_j$, it suffices that:

   $$\frac{a_j(G_j)}{v_0 + 2\gamma^2 a_j(G_j)^2 v_z} - \frac{a_j(G_j)}{v_0 + 2\gamma^2 a_j(G_j)^2 v_z}$$

   $$\propto a_j(G_j) \left[ v_0 + 2\gamma^2 a_j(G_j)^2 v_z \right] - a_j(G_j) \left[ v_0 + 2\gamma^2 a_j(G_j)^2 v_z \right]$$

   $$= \left[ a_j(G_j) - a_j(G_j) \right] \left[ v_0 - 2\gamma^2 a_j(G_j)^2 a_j(G_j) v_z \right]$$

   $$\propto 2\gamma^2 a_j(G_j)^2 a_j(G_j) v_z - v_0$$

   $$< 2\gamma^2 v_0^2 v_z - v_0$$

   $$< 0,$$

   where the second “$\propto$” follows from $a_j(G_j) < a_j(G_j)$, the first inequality follows from $a_j(G_j), a_j(G_j) < v_0$ (see Corollary 1) and the last inequality is obtained from Condition (3).

   If $G_j/\gamma^2 > v_0$, then $a_j(G_j) = 1/((\rho)/(v_{\theta_j})+(1-\rho)/(v_0-G_j/\gamma^2)) > 0$ so that $\lambda_j > 0$.

2. Consider the above expression for $\lambda_j$. It follows after taking partial derivatives that $\lambda_j$ increases in $\text{Var}(\tilde{R}_M)$. $\lambda_j$ also increases in $a_j(G_j)$ because:

   $$\frac{\partial \lambda_j}{\partial a_j(G_j)} \propto v_0 - 2\gamma^2 a_j(G_j)^2 v_z > v_0 - 2\gamma v_0^2 v_z > 0,$$

   where the first inequality follows from $a_j(G_j) < v_{\theta_j}$ (see Corollary 1), and the last inequality is obtained from Condition (3).

   Note that $\text{Var}(\tilde{R}_M)$ and $a_j(G_j)$ increase in $\rho$ (see Corollaries 1 and 2). It follows that $\lambda_j$ increases in $\rho$. ■
Proof of P3:

(1) From Corollary 1, \(a_j(G_j) < v_{\theta_j}\). Therefore, the \(\beta\)-adjusted expected return of the \(j\)th basic security is 
\[-\gamma [v_{\theta_j} + 2\gamma^2 a_j(G_j)^2 v_2 - a_2(G_j)]^2 \leq 0.\]

(2) From Corollary 1, \(a_j(G_j) < a_j(G_j)\) because \(G_j > G_j\). The difference between the \(\beta\)-adjusted expected returns of the \(j\)th and \(j'\)th basic securities is:
\[-\gamma [a_j(G_j) - a_j(G_j) - 2\gamma^2 \left(a_j(G_j)^2 - a_j(G_j)^2\right) v_2 - a_2(G_j)]^2 x_j\]
\[\propto 2\gamma^2 (a_j(G_j) + a_j(G_j)) v_2 - 1\]
\[< 4\gamma^2 v_{\theta_j} v_2 - 1\]
\[< 0,\]
where the “\(\propto\)” follows from \(a_j(G_j) < a_j(G_j)\), the first inequality follows from \(a_j(G_j)\), \(a_j(G_j) < v_{\theta_j}\) (see Corollary 1) and the last inequality is obtained from Condition (3).

(3) It follows from the above proof for part (ii) that the \(\beta\)-adjusted expected return of the \(j\)th basic security, 
\[-\gamma [v_{\theta_j} + 2\gamma^2 a_j(G_j)^2 v_2 - a_2(G_j)]^2 x_j,\]
increases in \(a_j(G_j)\). Also, \(a_j(G_j)\) increases in \(\rho\) (see Corollary 1). Therefore, this \(\beta\)-adjusted expected return increases in \(\rho\).

Proof of Corollary 3: Write Equations (10) and (11) as:
\[X_{NG_j}(P_j) - (\overline{\xi}_j + \overline{\xi}_j) \sim N\left(A_{NG}\overline{\xi}_j, A_{NG}^2 v_2\right),\]
\[X_{G_j}(P_j) - (\overline{\xi}_j + \overline{\xi}_j) \sim N\left(A_{G}\overline{\xi}_j, A_{G}^2 v_2\right),\]
where:
\[A_{NG} = \frac{a_j(G_j)}{v_{\theta_j}} - 1, \quad \text{and} \quad A_{G} = \frac{a_j(G_j)}{v_{\theta_j} - G_j / \gamma^2} - 1.\]

Here are some intermediate results we will use in the proof of this corollary. First, \(A_{NG} < 0\) and decreases in \(G_j\) from Corollary 1. Second, \(A_{G} = a_j(G_j) / (v_{\theta_j} - G_j / \gamma^2) - 1 = (1) / (\rho / v_{\theta_j}) (v_{\theta_j} - G_j / \gamma^2) + 1 - \rho \rangle > 0\), and increases in \(G_j\).

Using the functional form in Footnote 21 and the above new expressions for the distributions of \(X_{NG_j}(P_j) - (\overline{\xi}_j + \overline{\xi}_j)\) and \(X_{G_j}(P_j) - (\overline{\xi}_j + \overline{\xi}_j)[21]\), we can express the total expected trading volume in the \(j\)th basic security (Equation (12)) as:
\[T_j = 0.5 \rho E \left[|X_{NG_j}(P_j) - (\overline{\xi}_j + \overline{\xi}_j)|\right] + 0.5(1-\rho) E \left[|X_{G_j}(P_j) - (\overline{\xi}_j + \overline{\xi}_j)|\right]\]
\[= 0.5 \rho \left(-A_{NG} \sqrt{v_2}\right) \left[2 \phi \left(-\frac{\overline{\xi}_j}{\sqrt{v_2}}\right) - \frac{\overline{\xi}_j}{\sqrt{v_2}} \left(1 - 2 \Phi \left(\frac{\overline{\xi}_j}{\sqrt{v_2}}\right)\right)\right]\]
\[+ 0.5(1-\rho) A_{G} \sqrt{v_2}\left[2 \phi \left(-\frac{\overline{\xi}_j}{\sqrt{v_2}}\right) + \frac{\overline{\xi}_j}{\sqrt{v_2}} \left(1 - 2 \Phi \left(-\frac{\overline{\xi}_j}{\sqrt{v_2}}\right)\right)\right]\]
\[= [-\rho A_{NG} + (1-\rho) A_{G}] \times 0.5 \sqrt{v_2}\left[2 \phi \left(-\frac{\overline{\xi}_j}{\sqrt{v_2}}\right) - \frac{\overline{\xi}_j}{\sqrt{v_2}} \left(1 - 2 \Phi \left(\frac{\overline{\xi}_j}{\sqrt{v_2}}\right)\right)\right].\]

Footnote 21 indicates that the values in the brackets are positive. From the above analysis, \(A_{NG}\) decreases in \(G_j\), and \(A_{G}\) increases in \(G_j\). Therefore, \(T_j\) increases in \(G_j\).
RBF
11,2

Proof of Lemma 1: Consider Part (i) first. For $E(\tilde{R}_j) = \gamma a_j(G_j)\tilde{\gamma}_j$ to decrease in $\nu_0$, it is sufficient to show that $a_j(G_j)$ decreases in $\nu_0$. It follows from $PI$ and the constructed $G(\nu_0)$ in Equation (13) that:

$$a_j(G_j) = a_j(G(\nu_0)) \rightarrow \frac{1}{\nu_0} - \frac{1 - \rho}{\nu_0 - \frac{\nu_0}{\rho}(\nu_0 - \nu_0)}$$

(A1)

It follows that:

$$\frac{d\gamma a_j(G_j)}{d\nu_0} \rightarrow a_j(G_j)^2 \ell(\nu_0),$$

where:

$$\ell(\nu_0) \equiv \frac{\rho}{\nu_0^2} - \frac{1 - \rho}{\nu_0 - \frac{\nu_0}{\rho}(\nu_0 - \nu_0)}$$

(A2)

It is evident after taking derivatives that $\ell(\nu_0)$ decreases in $\nu_0$. Further $\ell(\nu_0) < 0$, and:

$$\ell(\nu_0) \propto \rho \frac{\nu_0 - \frac{\nu_0}{\rho}(\nu_0 - \nu_0)}{\nu_0 - \nu_0} \propto \rho \nu_0 - \nu_0 < 0$$

under the assumption $\rho < \min((\nu_0)/\nu_0, 1 - ((\nu_0)/\nu_0))$ (see Condition (14)). Taken together, $\ell(\nu_0) < 0$ and $a_j(G_j)$ decreases in $\nu_0$.

For Part (ii), it suffices to demonstrate that the basic security with a higher $\nu_0$ also has a higher $Cov(\tilde{R}_j, \tilde{R}_M)$. It follows from Equation (7) and the above analysis that:

$$\frac{dCov(\tilde{R}_j, \tilde{R}_M)}{d\nu_0} \rightarrow \left[1 + 4\gamma^2 a_j(G_j)\nu_0 a_j(G_j)^2 \ell(\nu_0)\right] \tilde{\gamma}_j$$

$$\propto 1 + 4\gamma^2 a_j(G_j)\nu_0 a_j(G_j)^2 \ell(\nu_0)$$

$$> 1 - 4\gamma^2 a_j(G_j)\nu_0 a_j(G_j)^2 \ell(\nu_0) \frac{1 - \rho}{\nu_0 - \frac{\nu_0}{\rho}(\nu_0 - \nu_0)}$$

$$> 1 - a_j(G_j)^2 \frac{1 - \rho}{\nu_0 - \frac{\nu_0}{\rho}(\nu_0 - \nu_0)}$$

$$> 1 - \frac{1}{1 - \rho} \frac{\nu_0}{\nu_0 - \nu_0}$$

where the first inequality follows from the expression of $\ell(\nu_0)$ (see Equation (A2)), the second inequality follows from $a_j(G_j) < \nu_0$ (see Corollary 1) and the assumption $\nu_0 < (1)/(\gamma^2 \nu_0) \min (1/4, \rho/2)$ (see Condition (3)), and the last inequality is obtained under the assumption $\rho < \min((\nu_0)/\nu_0, 1 - ((\nu_0)/\nu_0))$ (see Condition (14)).

Proof of Lemma 2: From Proposition 1, $(\text{Var}(\tilde{R}_j))/(\text{Var}(\tilde{R}_M)) = (\nu_0 + \gamma^2 a_j(G_j)^2 \nu_0)/(\text{Var}(\tilde{R}_M))$. This implies that “typical” basic securities with small $(\text{Var}(\tilde{R}_j))/(\text{Var}(\tilde{R}_M))$ also have small $(\nu_0)/(\text{Var}(\tilde{R}_M))$ and $(\gamma^2 a_j(G_j)^2 \nu_0)/(\text{Var}(\tilde{R}_M))$. We will use this property in the proof of this Lemma.
From Equation (15), Proposition 1, and our computation of \( \text{Cov}(\bar{R}_j, \bar{R}_M) \) in Appendix 2:

\[
\text{IVOL}_j = \text{Var} \left( \bar{R}_j \right) - \frac{\text{Cov} \left( \bar{R}_j, \bar{R}_M \right)}{\text{Var} \left( \bar{R}_M \right)} \cdot \text{IVOL}_j
\]

\[
= v_\theta + \gamma^2 a_j(G_j)^2 v_\nu - \frac{v_\theta + 2\gamma^2 a_j(G_j)^2 v_\nu}{\text{Var} \left( \bar{R}_M \right)} \cdot \text{ivol}_j
\]

It follows that for the \( j \)th and \( j \)'th typical basic securities:

\[
\text{IVOL}_j - \text{IVOL}_j' = \left[ v_\theta + \gamma^2 a_j(G(v_\theta))^2 v_\nu \right] - \left[ v_\theta + \gamma^2 a_j(G(v_\theta'))^2 v_\nu \right]
\]

\[
- \frac{v_\theta + 2\gamma^2 a_j(G(v_\theta))^2 v_\nu + v_\theta + 2\gamma^2 a_j(G(v_\theta'))^2 v_\nu}{\text{Var} \left( \bar{R}_M \right)} \cdot \text{IVOL}_j
\]

\[
\approx v_\theta - v_\theta' + \gamma^2 v_\nu \left[ a_j(G(v_\theta))^2 - a_j(G(v_\theta'))^2 \right],
\]

where \( \delta = ((v_\theta + 2\gamma^2 a_j(G(v_\theta))^2 v_\nu + v_\theta + 2\gamma^2 a_j(G(v_\theta'))^2 v_\nu) / (\text{Var}(\bar{R}_M))) \), and the last “≈” follows from the above derived property that for the \( j \)th and \( j \)'th typical basic securities, \( \delta \) is small. For \( \text{IVOL}_j > \text{IVOL}_j' \), it suffices that:

\[
\frac{d}{v_\theta} \left[ v_\theta + \gamma^2 v_\nu a_j(G_j)^2 \right] \rightarrow 1 + 2\gamma^2 a_j(G_j) v_\nu a_j(G_j)^2 \ell(v_\theta)
\]

\[
> 1 - 2\gamma^2 a_j(G_j) v_\nu a_j(G_j)^2 \frac{1 - \rho}{(v_\theta - \frac{\gamma_\theta}{\gamma_\nu} (v_\theta - v_\nu))^2 v_\theta - v_\nu}
\]

\[
> 1 - a_j(G_j)^2 \frac{1 - \rho}{(v_\theta - \frac{\gamma_\theta}{\gamma_\nu} (v_\theta - v_\nu))^2 v_\theta - v_\nu}
\]

\[
\rightarrow -1 - \frac{1 - \rho}{\left( \frac{v_\theta}{v_\nu} - \frac{\gamma_\theta}{\gamma_\nu} (v_\theta - v_\nu) \right)^2 v_\theta - v_\nu}
\]

\[
+ \frac{1 - \rho}{\left( \frac{v_\theta}{v_\nu} - \frac{\gamma_\theta}{\gamma_\nu} (v_\theta - v_\nu) \right)^2 v_\theta - v_\nu}
\]

\[
\rightarrow -1 - \frac{1 - \rho}{\left( \frac{v_\theta}{v_\nu} - \frac{\gamma_\theta}{\gamma_\nu} (v_\theta - v_\nu) \right)^2 v_\theta - v_\nu}.
\]
where the first inequality follows from the expression of $\ell(v_\theta)$ (see Equation (A2)), the second inequality follows from $a_j(G_j) < v_{\theta_j}$ (see Corollary 1) and the assumption $v_{\theta_j} < (1/\gamma^2 v_\theta) \min(1/4, \rho/2)$ (see Condition (3)), and the last inequality obtains under the assumption $\rho < \min((v_\theta/v_\theta), 1 - ((v_{\theta_j})/(v_{\theta_j} - v_\theta)))$ (see Condition (14)).

Proof of P4: Taking derivatives of the $\beta$-adjusted expected return, $-\gamma(v_{\theta_j} + 2\gamma^2 a_j(G_j)^2 v_{\theta_j} - a_j(G_j))\varepsilon_j$, wrt $v_{\theta_j}$ yields:

$$
\frac{d}{dv_{\theta_j}} \left[-\gamma(v_{\theta_j} + 2\gamma^2 a_j(G_j)^2 v_{\theta_j} - a_j(G_j))\varepsilon_j \right] = -1 + \left[1 - 4\gamma^2 a_j(G_j) v_{\theta_j} \right] a_j(G_j)^2 \ell(v_{\theta_j}).
$$

It follows from $a_j(G_j) < v_{\theta_j}$ (see Corollary 1) and the assumption $v_{\theta_j} < (1/(\gamma^2 v_\theta)) \min(1/4, \rho/2)$ (see Condition (3)) that $1 - 4\gamma^2 a_j(G_j) v_{\theta_j} > 0$. Also, the proof of Lemma 1 shows that under the assumption $\rho < \min((v_\theta/v_{\theta_j}), 1 - ((v_{\theta_j})/(v_{\theta_j} - v_\theta)))$ (see Condition (14)), $\ell(v_{\theta_j}) < 0$. Therefore, the above derivative is negative, and the $\beta$-adjusted expected return decreases in $v_{\theta_j}$.

Proof of P5: It follows from Equation (16) that:

$$
E(\tilde{R}_M) = \sum_{j=1}^{N+k} E\left[\tilde{\varepsilon}_j, \tilde{\varepsilon}_j + \gamma a_j(G_j) \left(\tilde{\varepsilon}_j^2 + 2\tilde{\varepsilon}_j + \tilde{\varepsilon}_j^2\right)\right] = \sum_{j=1}^{N+k} \left[\gamma a_j(G_j) \left(\tilde{\varepsilon}_j^2 + \tilde{\varepsilon}_j + \tilde{\varepsilon}_j^2\right)\right] > 0.
$$

Proof of Lemma 3: Consider the security payoffs in Equation (2). The $N + k$th $(k = 1, \ldots, K)$ original security has a payoff:

$$
V_{N+k} = \tilde{f}_k.
$$

Write its price as $P(V_{N+k}) = P_{N+k}$. The $j$th $(j = 1, \ldots, N)$ original security takes the linear form, i.e.:

$$
V_j = \nabla_j + \sum_{k=1}^{K} \beta_{jk} \tilde{f}_k + \tilde{c}_j.
$$

Write its price as:

$$
P(V_j) = \nabla_j + \sum_{k=1}^{K} \beta_{jk} P_{N+k} + P_j.
$$

The demands of the $i$th type-$D$ non-$G$ or $G$ trader for the original securities are denoted $y_i (j = 1, \ldots, N)$ and $y_{N+k} (k = 1, \ldots, K)$. His/her wealth at Date 1 can be expressed as:

$$
W_{i1} = W_{i0} + \sum_{j=1}^{N} y_j (V_j - P(V_j)) + \sum_{k=1}^{K} y_{N+k} (V_{N+k} - P(V_{N+k}))
$$

$$
= W_{i0} + \sum_{j=1}^{N} y_j \left[ \sum_{k=1}^{K} \beta_{jk} \left( \tilde{f}_k - P_{N+k} \right) + \tilde{c}_j - P_j \right] + \sum_{k=1}^{K} y_{N+k} \left( \tilde{f}_k - P_{N+k} \right)
$$

$$
= W_{i0} + \sum_{j=1}^{N} y_j (\tilde{c}_j - P_j) + \sum_{k=1}^{K} y_{N+k} \left[ \sum_{j=1}^{N} y_j \beta_{jk} \right] \left( \tilde{f}_k - P_{N+k} \right) \quad \text{(A3)}
$$
\[ W_0 + \sum_{j=1}^{N} Y_j (\ell_j - P_j) + \sum_{k=1}^{K} Y_{N+k} (\ell_{k} - P_{N+k}), \] 

\( (A4) \)

where the last equality is obtained from \( Y_j = y_j \) \((j = 1, \ldots, N)\) and \( Y_{N+k} = y_{N+k} + \sum_{j=1}^{N} y_j \beta_{jk} \) \( (k = 1, \ldots, K).\)

The expression in Equation (A3) indicates that the trader can choose the demands for the original securities indicated by \( y_j \) \((j = 1, \ldots, N)\) and \( Y_{N+k} \) \((k = 1, \ldots, K)\), conditional on the prices of the original securities \( P(V) \) and \( P(V_{N+k}) \). The expression in Equation (A4) indicates that the trader can equivalently choose the demands for the basic securities indicated by \( Y_j \) \((j = 1, \ldots, N)\) and \( Y_{N+k} \) \((k = 1, \ldots, K)\), conditional on the prices of the basic securities \( P_j \) and \( P_{N+k} \). Now, the market clearing condition for the basic securities is implied by the market clearing condition for the original security; that is, \( \forall j = 1, \ldots, N: \)

\[ \rho Y_{NG,j} + (1-\rho) Y_{G,j} = \rho y_{NG,j} + (1-\rho)y_{G,j} = s_j = \xi_j + \eta_j, \]

and \( \forall k = 1, \ldots, K: \)

\[ \rho Y_{NG,k} + (1-\rho) Y_{G,k} = \rho y_{NG,k} + (1-\rho)y_{G,k} = s_k = \xi_k + \eta_k. \]

There are two additional observations. First, \( G \) traders’ utility from trading emanates from \( Y_j \) \((j = 1, \ldots, N)\) net holdings of \( \ell_j \) and the \( Y_{N+k} \) \((k = 1, \ldots, K)\) net holdings of \( f_{k} \), respectively. Second, we can use the same analysis as in the paper for the basic securities to find \( P_j \) and \( P_{N+k} \), which are identical to those presented in PI.

Proof of P6:

1. From Equations (21) and (23):

\[ E(R^{ANG}_j) = \gamma v_0 \xi_j > E(R^{AG}_j) = \gamma (v_0 - G_j / \gamma^2) \xi_j, \]

\[ \text{Var}(R^{ANG}_j) = v_0 + \gamma^2 v_0^2 v_2 > \text{Var}(R^{AG}_j) = v_0 + \gamma^2 (v_0 - G_j / \gamma^2)^2 v_2. \]

2. From Equation (21), it follows that the return of the market portfolio in the all-non-\( G \) economy is given by:

\[ R^{ANG}_M = \sum_{j=1}^{N+K} (\xi_j + \eta_j) \]

\[ R^{ANG}_j = \sum_{j=1}^{N+K} [(\xi_j + \eta_j) (\ell_j + \gamma v_0 (\xi_j + \eta_j))] \]

It follows immediately that:

\[ E(R^{ANG}_M) = \sum_{j=1}^{N+K} E[(\xi_j + \eta_j) (\ell_j + \gamma v_0 (\xi_j + \eta_j))] = \sum_{j=1}^{N+K} \left[ \gamma v_0 (\xi_j^2 + v_2) \right], \]

\[ \text{Var}(R^{ANG}_M) = \sum_{j=1}^{N+K} \left[ (v_0 + 4 \gamma^2 v_0^2 v_2) \xi_j^2 + v_0 v_2 + 2 \gamma^2 v_0^2 v_2^2 \right]. \]
where the last equality follows from Equation (6) because the all-non-$G$ economy can be viewed as the case in which $a(G_j) = v_0$. From Equation (23), it follows that the return of the market portfolio in the all-$G$ economy is given by:

$$R^A_G = \sum_{j=1}^{N+K} (\bar{z}_j + \bar{z}_j) \hat{R}^G_j = \sum_{j=1}^{N+K} \left[ (\bar{z}_j + \bar{z}_j) (\bar{\theta}_j + \gamma (v_0 - G_j / \gamma^2) (\bar{z}_j + \bar{z}_j)) \right].$$

It follows immediately that:

$$E\left(R^A_G\right) = \sum_{j=1}^{N+K} E\left[ (\bar{z}_j + \bar{z}_j) (\bar{\theta}_j + \gamma (v_0 - G_j / \gamma^2) (\bar{z}_j + \bar{z}_j)) \right] = \sum_{j=1}^{N+K} \left[ \gamma (v_0 - G_j / \gamma^2) (\bar{z}_j^2 + v_0) \right],$$

$$\text{Var}\left(R^A_G\right) = \sum_{j=1}^{N+K} \left[ \left( v_0 + 4\gamma^2 (v_0 - G_j / \gamma^2)^2 v_z \right) \bar{z}_j^2 + v_0 v_z + 2\gamma^2 (v_0 - G_j / \gamma^2)^2 v_z^2 \right],$$

where the last equality follows from Equation (6) because the all-$G$ economy is the case in which $a(G_j) = v_0 - G_j / \gamma^2$.

A direct comparison indicates that $E\left(R^A_{M,NG}\right) > E\left(R^A_M\right)$ and $\text{Var}\left(R^A_{M,NG}\right) > \text{Var}\left(R^A_M\right)$:

(3) Note that $\lambda^A_{M,NG} = \langle (\gamma v_0 \text{Var}(R^A_M))/(v_0 + 2\gamma^2 v_0^2 v_z) \rangle$ and $\lambda^A_M = \langle (\gamma(v_0 - G_j / \gamma^2)\text{Var}(R^A_G))/(v_0 + 2\gamma^2 (v_0 - G_j / \gamma^2)^2 v_z) \rangle$. From Part (ii) $\text{Var}(R^A_M) > \text{Var}(R^A_G)$. Thus, for $\lambda^A_{M,NG} > \lambda^A_M$, it suffices that:

$$\frac{v_0}{v_0 + 2\gamma^2 v_0^2 v_z} - \frac{v_0 - G_j / \gamma^2}{v_0 + 2\gamma^2 (v_0 - G_j / \gamma^2)^2 v_z} \propto v_0 \left[ v_0 + 2\gamma^2 (v_0 - G_j / \gamma^2) v_z \right] - (v_0 - G_j / \gamma^2) \left( v_0 + 2\gamma^2 v_0^2 v_z \right)$$

$$\propto v_0 - 2\gamma^2 (v_0 - G_j / \gamma^2) v_0 v_z$$

$$> v_0 - 2\gamma^2 v_0^2 v_z$$

$$> 0,$$

where the last inequality is obtained from Condition (3).

Proof of $P7$. Note that $\lambda_j = \langle \gamma a_j(G_j) \text{Var}(R_M) \rangle/(v_0 + 2\gamma^2 a_j(G_j)^2 v_z)$ and $\lambda^j_j = \langle (\gamma / \rho v_0) \text{Var}(R_M) \rangle/(v_0 + 2(\gamma / \rho)^2 v_0^2 v_z)$.

For $\lambda < \lambda^j_j$, it suffices to show that:

$$\frac{\gamma a_j(G_j)}{v_0 + 2\gamma^2 a_j(G_j)^2 v_z} - \frac{(\gamma / \rho) v_0}{v_0 + 2(\gamma / \rho)^2 v_0^2 v_z} \propto \gamma a_j(G_j) \left[ v_0 + 2(\gamma / \rho)^2 v_0 v_z \right] - (\gamma / \rho) v_0 \left[ v_0 + 2\gamma^2 a_j(G_j)^2 v_z \right]$$

$$= \left[ \gamma a_j(G_j) - (\gamma / \rho) v_0 \right] \left[ v_0 - 2\gamma a_j(G_j) (\gamma / \rho) v_0 v_z \right]$$

$$\propto 2\gamma^2 a_j(G_j) v_z - \rho$$

$$< 2\gamma^2 v_0 v_z - \rho$$

$$> 0,$$

where the second “$\propto$” and the first inequality follow from $a_j(G_j) < v_0$ (see Corollary 1), and the last inequality is obtained from Condition (3).
Proof of P8: Taking the \textit{ex ante} expectation (over }\bar{z}\textit{) of Equation (29) yields:

\[
E^{\text{ex}}[U_{G^4}(W_t, X_t)] = E \left[ -\exp \left[ -\gamma M_1 - 0.5 \frac{\gamma^2 a(G, \rho_1) z (\zeta_1 + \bar{z})^2}{v - G/\gamma^2} \right] \right].
\]

It is straightforward to show, after taking derivatives, that \(E^{\text{ex}}[U_{G^4}(W_t, X_t)]\) increases in \(a(G, \rho_1)\)  and \(a(G, \rho_1)\) decreases in \(\rho_1\). Therefore, \(E^{\text{ex}}[U_{G^4}(W_t, X_t)]\) decreases in \(\rho_1\). It is obvious that \(E[U_{G^4}(W_t, Y)]\) in Equation (30) increases in \(G^4\).

A \(G^4\) trader’s decision to participate in the stock market, instead of gambling through the alternative avenue, is based on:

\[
\zeta(\rho_1, G^4) \equiv E^{\text{ex}}[U_{G^4}(W_t, X_t)] - E[U_{G^4}(W_t, Y)].
\]

If \(\zeta(\rho_1, G^4) > 0\), then he/she will participate the stock market. Otherwise, he/she will not.

It follows from the above analysis that \(\zeta(\rho_1, G^4)\) decreases in \(\rho_1\) and \(G^4\). Therefore, the unique \(\rho_1 \in [0, \rho^A]\) is determined as follows:

- if \(\zeta(0, G^4) \leq 0\), then \(\rho_1 = 0\);
- if \(\zeta(\rho^A, G^4) \geq 0\), then \(\rho_1 = \rho^A\); and
- if \(\zeta(\rho^A, G^4) < 0 < \zeta(0, G^4)\), then an interior \(\rho_1\) is specified by \(\zeta(\rho_1, G^4) = 0\).

Further, \(\rho_1\) decreases in \(G^4\).

Proof of Corollary 4: It follows from Equations (25)–(28) that:

\[
X_{NG}(P) - (\bar{z}_1 + \bar{z}) \sim N \left( A_{NG}(\rho_1) \frac{z}{\sqrt{v}} , A_{NG}(\rho_1)^2 v_z \right),
\]

\[
X_{G}(P) - (\bar{z}_1 + \bar{z}) \sim N \left( A_{G}(\rho_1) \frac{z}{\sqrt{v}} , A_{G}(\rho_1)^2 v_z \right),
\]

\[
X_{G^4}(P) \sim N \left( \frac{a(G, \rho_1)}{v - G/\gamma^2} \frac{\zeta}{\sqrt{v_z}} , \left( \frac{a(G, \rho_1)}{v - G/\gamma^2} \right)^2 v_z \right),
\]

where:

\[
A_{NG}(\rho_1) \equiv \frac{a(G, \rho_1)}{v} - 1, \quad \text{and} \quad A_{G}(\rho_1) \equiv \frac{a(G, \rho_1)}{v - G/\gamma^2} - 1.
\]

If \(A_G(\rho_1) < 0\), then, using the fact listed in Footnote 21, we can express the total expected trading volume (Equation (31)) as:

\[
T(\rho_1) = 0.5 \rho \left[ -A_{NG}(\rho_1) \sqrt{v_z} \left( 2\phi \left( -\frac{\zeta}{\sqrt{v_z}} \right) - \frac{\zeta}{\sqrt{v_z}} \left( 1 - 2\Phi \left( \frac{\zeta}{\sqrt{v_z}} \right) \right) \right) \right] + 0.5(1-\rho) \left[ -A_{G}(\rho_1) \sqrt{v_z} \left( 2\phi \left( -\frac{\zeta}{\sqrt{v_z}} \right) - \frac{\zeta}{\sqrt{v_z}} \left( 1 - 2\Phi \left( \frac{\zeta}{\sqrt{v_z}} \right) \right) \right) \right] + 0.5 \rho a(G, \rho_1) \sqrt{v_z} \left( 2\phi \left( \frac{\zeta}{\sqrt{v_z}} \right) + \frac{\zeta}{\sqrt{v_z}} \left( 1 - 2\Phi \left( \frac{\zeta}{\sqrt{v_z}} \right) \right) \right) + 0.5 \rho a(G, \rho_1) \sqrt{v_z} \left( 2\phi \left( -\frac{\zeta}{\sqrt{v_z}} \right) - \frac{\zeta}{\sqrt{v_z}} \left( 1 - 2\Phi \left( \frac{\zeta}{\sqrt{v_z}} \right) \right) \right) = -\rho A_{NG}(\rho_1)(1-\rho)A_{G}(\rho_1) + \rho_1 a(G, \rho_1) \left( \frac{\zeta}{\sqrt{v_z}} \right)\right] + \rho_1 a(G, \rho_1) \left( \frac{\zeta}{\sqrt{v_z}} \right)\right] \times 0.5 \sqrt{v_z} \left( 2\phi \left( -\frac{\zeta}{\sqrt{v_z}} \right) - \frac{\zeta}{\sqrt{v_z}} \left( 1 - 2\Phi \left( \frac{\zeta}{\sqrt{v_z}} \right) \right) \right) \right] \right].
Taking derivative wrt $\rho_1$ yields:

$$dT(\rho_1) \propto \frac{d}{d\rho_1} \left[ -\rho A_{NG}(\rho_1) - (1-\rho)A_G(\rho_1) + \rho_1 a(G,\rho_1) \right]$$

$$= \frac{a(G,\rho_1)}{v - G/\gamma^2} + \left( -\frac{\rho}{v} \frac{1-\rho}{v - G/\gamma^2} + \frac{\rho_1}{v - G/\gamma^2} \right) a(G,\rho_1)$$

$$\propto 1 - \left( \frac{\rho}{v} \frac{1-\rho}{v - G/\gamma^2} + \frac{\rho_1}{v - G/\gamma^2} \right) a(G,\rho_1)$$

$$= 1 - \frac{\rho}{v} \frac{1-\rho}{v - G/\gamma^2} + \frac{\rho_1}{v - G/\gamma^2} > 0.$$

If $A_{NG} > 0$, then we can use a similar analysis to show that $dT(\rho_1)/d\rho_1 > 0$. Taken together, $dT(\rho_1)/d\rho_1 > 0$.

Note that $\rho_1$ decreases in $G^4$ (see P8). It follows that $T(\rho_1)$ decreases in $G^4$.

Proof of P9: We solve for the equilibrium and prove the proposition using backward induction, in three steps.

- Step 1: in this step, suppose $\tilde{\theta}_1 > H_\theta$ and therefore $P_1 > P_0$ (which we will show below). An agent remains a non-$G$ trader (becomes a $G$ trader) with probability $\rho (1-\rho)$. Thus, there is a mass $\rho$ of non-$G$ traders and a mass $1-\rho$ of $G$ traders at Date 2.

Focus on Date 2 for the moment. Write the $i$th $G$ trader’s wealth at Date 3 as $W_i = W_{i2} + X_i(V-P_2)$. His/her expected utility at Date 2 can be expressed as:

$$E[U_{NG}(W_{i3})|\tilde{\theta}_1, \tilde{\theta}_2] = -\exp[-\gamma W_{i2}]$$

$$-\gamma \left[ X_{i2}(E(V|\tilde{\theta}_1, \tilde{\theta}_2)-P_2) - 0.5\gamma \text{Var}(V|\tilde{\theta}_1, \tilde{\theta}_2) \right].$$

He/she needs to choose $X_{i2}$ to maximize this expected utility. The foc implies that his/her demand can be expressed as:

$$X_{iNG2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2) = \frac{E(V|\tilde{\theta}_1, \tilde{\theta}_2)-P_2}{\gamma \text{Var}(V|\tilde{\theta}_1, \tilde{\theta}_2)} = \frac{\nabla + \tilde{\theta}_1 + \tilde{\theta}_2 - P_2}{\gamma v_0}. \quad (A6)$$

An $i$th $G$ trader’s expected utility at Date 2 can be expressed as:

$$E[U_G(W_{i3}, X_{i2})|\tilde{\theta}_1, \tilde{\theta}_2] = -\exp[-\gamma W_{i2}]$$

$$-\gamma \left[ X_{i2}(E(V|\tilde{\theta}_1, \tilde{\theta}_2)-P_2) - 0.5\gamma X^2_{i2} \text{Var}(V|\tilde{\theta}_1, \tilde{\theta}_2) + 0.5GX^2_{i2}/\gamma \right].$$

He/she needs to choose $X_{i2}$ to maximize this expected utility. The foc implies that his demand can be expressed as:

$$X_{iG2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2) = \frac{E(V|\tilde{\theta}_1, \tilde{\theta}_2)-P_2}{\gamma \text{Var}(V|\tilde{\theta}_1, \tilde{\theta}_2)-G/\gamma} = \frac{\nabla + \tilde{\theta}_1 + \tilde{\theta}_2 - P_2}{\gamma (v_0 - G/\gamma^2)}. \quad (A8)$$

It follows from Equations (A6) and (A8) that the market clearing condition requires:

$$\xi = \rho \cdot X_{iNG2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2) + (1-\rho) \cdot X_{iG2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2)$$

$$= \frac{\rho}{\gamma v_0} (\nabla + \tilde{\theta}_1 + \tilde{\theta}_2 - P_2) + \frac{1-\rho}{\gamma (v_0 - G/\gamma^2)} (\nabla + \tilde{\theta}_1 + \tilde{\theta}_2 - P_2).$$
Therefore:

\[ P_2 = V + \hat{\theta}_1 + \hat{\theta}_2 - \gamma a(G) \xi, \]  \hspace{1cm} (A9)

where:

\[ a(G) = \frac{1}{\frac{\gamma}{v_0} + \frac{1}{v_0 - G/\gamma^2}}. \]

Consider a \( G \) trader’s expected utility at Date 2 in Equation (A5). Plugging in the optimal demand for the risky security from Equation (A6) yields:

\[
E[U_{NG}(W_{32})|\tilde{\theta}_1, \tilde{\theta}_2] = -\exp \left[-\gamma W_{22} - 0.5 \left( \frac{V + \tilde{\theta}_1 + \tilde{\theta}_2 - P_2}{v_0} \right)^2 \right] 
\]

\[ = -\exp \left[-\gamma W_{22} - 0.5 \frac{(\gamma a(G) \xi)^2}{v_0} \right]. \hspace{1cm} (A10) \]

Consider a \( G \) trader’s expected utility at Date 2 in Equation (A7). Plugging in the optimal demand for the risky security from Equation (A8) yields:

\[
E[U_G(W_{32}, X_{32})|\tilde{\theta}_1, \tilde{\theta}_2] = -\exp \left[-\gamma W_{22} - 0.5 \left( \frac{V + \tilde{\theta}_1 + \tilde{\theta}_2 - P_2}{v_0 - G/\gamma^2} \right)^2 \right] 
\]

\[ = -\exp \left[-\gamma W_{22} - 0.5 \frac{(\gamma a(G) \xi)^2}{v_0 - G/\gamma^2} \right]. \hspace{1cm} (A11) \]

Now focus on Date 1. Write an \( i \) th trader’s wealth at Date 2 as \( W_{22} = W_{32} + X_{32}(P_2 - P_1) \). It follows from Equations (A10) and (A11) that his/her expected utility at Date 1 can be expressed as:

\[
E[U(W_{i3})|\tilde{\theta}_i] = \rho E[U_{NG}(W_{3i})|\tilde{\theta}_i] + (1-\rho) E[U_G(W_{3i}, X_{3i})|\tilde{\theta}_i] 
\]

\[ = -\exp \left[-\gamma W_{i1} - \gamma \left( X_{i1} (E(P_2|\tilde{\theta}_i) - P_1) - 0.5 \gamma X_{i1}^2 \text{Var}(P_2|\tilde{\theta}_i) \right) \right] 
\]

\[ \cdot \left[ \rho \exp \left( -0.5 \frac{(\gamma a(G) \xi)^2}{v_0} \right) + (1-\rho) \exp \left( -0.5 \frac{(\gamma a(G) \xi)^2}{v_0 - G/\gamma^2} \right) \right]. \hspace{1cm} (A12) \]

He/she needs to choose \( X_{i1} \) to maximize this expected utility. The fcc implies that his/her demand can be expressed as:

\[
X_{i1}(\tilde{\theta}_i, P_1) = \frac{E(P_2|\tilde{\theta}_i) - P_1}{\gamma \text{Var}(P_2|\tilde{\theta}_i)} = \frac{V + \tilde{\theta}_1 - \gamma a(G) \xi - P_1}{\gamma v_0}. \]

The market clearing condition, \( X_{i1}(\tilde{\theta}_i, P_1) = \xi \), implies

\[ P_1 = V + \tilde{\theta}_1 - \gamma v_0 \xi - \gamma a(G) \xi. \hspace{1cm} (A13) \]

Substituting the derived \( X_i(P_1, \tilde{\theta}_i) \) and \( P_1 \) back into his expected utility at Date 1 using Equation (A13) yields:

\[
E[U(W_{i3})|\tilde{\theta}_i] = -\exp \left[-\gamma W_{i1} - 0.5 \frac{(\gamma v_0 \xi)^2}{v_0} \right] 
\]

\[ \cdot \left[ \rho \exp \left( -0.5 \frac{(\gamma a(G) \xi)^2}{v_0} \right) + (1-\rho) \exp \left( -0.5 \frac{(\gamma a(G) \xi)^2}{v_0 - G/\gamma^2} \right) \right]. \hspace{1cm} (A14) \]
Step 2: in this step, suppose \( \tilde{\theta}_1 \leq H_\theta \) and therefore \( P_1 \leq P_0 \) (as we show). An agent remains a non-\( G \) trader. We can use a similar derivation as in Step 1, except that we impose \( G = 0 \) (note that \( a(0) = \nu_0 \)), to show:

\[
P_2 = V + \tilde{\theta}_1 + \gamma \nu_0 \xi,
\]

\[
P_1 = V + \tilde{\theta}_1 - 2\gamma \nu_0 \xi. \tag{A15}
\]

Moreover, the agent’s expected utility at Date 1 is given by:

\[
EU(W_{i2} | \tilde{\theta}_1) = -\exp \left[ -\gamma W_{i1} - \frac{(\gamma \nu_0 \xi)^2}{\nu_0} \right]. \tag{A16}
\]

Step 3: we now focus on Date 0. Since this date occurs before \( \tilde{\theta}_1 \) is released and \( P_1 \) is formed, all agents have identical preferences and beliefs and hold the same long position of the risky asset to clear the market.

Write an \( i \)th trader’s wealth at Date 1 as \( W_{i2} = W_{i0} + X_{i0}(P_1 - P_0) \). If \( \tilde{\theta}_1 > H_\theta \) so that \( P_1 > P_0 \) he/she may become a \( G \) trader. He will then be in the regime characterized by Equations (A13) and (A14). If \( \tilde{\theta}_1 \leq H_\theta \) so that \( P_1 \leq P_0 \), he/she will remain a non-\( G \) trader for sure. He/she will be in the regime characterized by Equations (A15) and (A16). Accounting for both cases, we can write his/her expected utility at date 0 as:

\[
E[U(W_{i2})] = -\int_{H_\theta}^{\infty} \exp \left[ -\gamma W_{i0} - \gamma X_{i0}(V + \tilde{\theta}_1 - \gamma \nu_0 \xi - \gamma a(G) \xi - P_0) - 0.5 \frac{(\gamma \nu_0 \xi)^2}{\nu_0} \right] \times \\
\cdot \rho \exp \left( -0.5 \frac{(\gamma a(G) \xi)^2}{\nu_0} \right) + (1 - \rho) \exp \left( -0.5 \frac{(\gamma a(G) \xi)^2}{\nu_0} \right) \frac{d\Phi \left( \frac{\tilde{\theta}_1}{\nu_0} \right)}{
u_0} \\
- \int_{-\infty}^{H_\theta} \exp \left[ -\gamma W_{i0} - \gamma X_{i0}(V + \tilde{\theta}_1 - 2\gamma \nu_0 \xi - P_0) - 0.5 \frac{(\gamma \nu_0 \xi)^2}{\nu_0} \right] d\Phi \left( \frac{\tilde{\theta}_1}{\nu_0} \right),
\]

where \( \Phi() \) is the cumulative density function of standard normal distribution. The agent needs to choose \( X_{i0} \) to maximize this expected utility. Taking the foc wrt \( X_{i0} \), imposing the market clearing condition \( X_{i0} = \xi \), and simplifying yields:

\[
0 = \int_{H_\theta}^{\infty} \frac{V + \tilde{\theta}_1 - \gamma \nu_0 \xi - \gamma a(G) \xi - P_0}{\exp(\gamma \xi \tilde{\theta}_1)} \times \\
\cdot \rho \exp \left( -0.5 \frac{(\gamma a(G) \xi)^2}{\nu_0} \right) + (1 - \rho) \exp \left( -0.5 \frac{(\gamma a(G) \xi)^2}{\nu_0} \right) \frac{d\Phi \left( \frac{\tilde{\theta}_1}{\nu_0} \right)}{\nu_0} \\
+ \int_{-\infty}^{H_\theta} \frac{V + \tilde{\theta}_1 - 2\gamma \nu_0 \xi - P_0}{\exp(\gamma \xi (\tilde{\theta}_1 - \gamma (\nu_0 - a(G)) \xi))} \times \exp \left( -0.5 \frac{(\gamma \nu_0 \xi)^2}{\nu_0} \right) \frac{d\Phi \left( \frac{\tilde{\theta}_1}{\nu_0} \right)}{\nu_0}. \tag{A17}
\]

If \( \tilde{\theta}_1 = H_\theta \), then \( P_1 = P_0 \), where \( P_1 \) is given by Equation (A15). This implies:

\[
P_0 = V + H_\theta - 2\gamma \nu_0 \xi.
\]
Substituting this expression for \( P_0 \) back into Equation (A17) yields:

\[
0 = \int_{H_0}^{\infty} \frac{\tilde{\theta}_1 - H_0 + \gamma(v_0 - a(G)\xi)}{\exp(\gamma^2 \tilde{\theta}_1)} \exp\left(-0.5\frac{(\gamma a(G)\xi)^2}{v_0 - G^2/\gamma^2}\right) d\Phi\left(\frac{\tilde{\theta}_1}{\sqrt{v_0}}\right)
\]

\[
\cdot \left[\rho \exp\left(-0.5\frac{(\gamma a(G)\xi)^2}{v_0}\right) + (1 - \rho)\exp\left(-0.5\frac{(\gamma a(G)\xi)^2}{v_0 - G^2/\gamma^2}\right)\right] d\Phi\left(\frac{\tilde{\theta}_1}{\sqrt{v_0}}\right).
\]

It is straightforward to show that the right-hand side of this equation decreases in \( H_0 \), and is positive (negative) if \( H_0 < -\infty (H_0 > \infty) \). Therefore, \( H_0 \) is uniquely determined by this equation.

Proof of Corollary 5:

(1) Note from P7 that regardless of whether \( P_1 > P_0 \) or \( P_1 \leq P_0 \), \( P_2 - P_1 = \tilde{\theta}_2 + \gamma v_0 \xi \). Therefore, Var \( (P_2 - P_1) | P_1 > P_0 \) = Var \( (P_2 - P_1) | P_1 \leq P_0 \). It follows from P9 that:

\[
V - P_2 = \begin{cases} 
\tilde{\theta}_3 + \gamma a(G)\xi & \text{if } P_1 > P_0, \\
\tilde{\theta}_3 + \gamma v_0 \xi & \text{if } P_1 \leq P_0.
\end{cases}
\]

We then have:

\[
\text{Var}(V - P_2 | P_1 > P_0) = v_0 + (\gamma a(G))^2 \text{Var}(\xi),
\]

\[
\text{Var}(V - P_2 | P_1 \leq P_0) = v_0 + (\gamma v_0)^2 \text{Var}(\xi).
\]

Note from P9 that \( 0 < a(G) < v_0 \). It follows immediately that \( \text{Var}(V - P_2 | P_1 > P_0) < \text{Var}(V - P_2 | P_1 \leq P_0) \).

(2) Note that before Date 2, all agents are identical so they have identical holdings, \( X_0 = X_1 = \xi \). If \( \tilde{\theta}_1 > H_0 \) so that \( P_1 > P_0 \), a mass \( \rho \) of agents become non-\( G \) traders and have the following demand (from Equations (A6) and (A9)):

\[
X_{NG,2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2) = \frac{\bar{V} + \tilde{\theta}_1 + \tilde{\theta}_2 - P_2}{\gamma v_0} = a(G)\xi.
\]

A mass \( 1 - \rho \) of agents become \( G \) traders and their demand is given by the following (from Equations (A8) and (A9)):

\[
X_{G,2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2) = \frac{a(G)\xi}{v_0 - G^2/\gamma^2}.
\]

We can express expected trading volume as:

\[
T_2(P_1 > P_0) = 0.5 \rho E \left[ X_{NG,2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2) - \xi \right] + 0.5(1 - \rho) E \left[ X_{G,2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2) - \xi \right]
= (1 - \rho) E \left[ X_{G,2}(\tilde{\theta}_1, \tilde{\theta}_2, P_2) - \xi \right]
= (1 - \rho) \left( \frac{a(G)}{v_0 - G^2/\gamma^2} - 1 \right) \xi,
\]

where the second equality can easily be verified. Now, if \( \tilde{\theta}_1 \leq H_0 \) so that \( P_1 \leq P_0 \), all agents will become non-\( G \) traders and continue to all hold the same amount of the stock, \( X_2 = \xi \). Thus, the expected trading volume can be written as \( T_2(P_1 > P_0) = T_2(P_1 \leq P_0) \). It follows that \( T_2(P_1 > P_0) > T_2(P_1 \leq P_0) \).
Appendix 2

Computation of market portfolio’s return volatility from the definition in Section 2.2

From \( P_I \), the returns of the \( j \)th basic security and the market portfolio are given by:

\[
\tilde{R}_j = \tilde{\theta}_j - P_j = \tilde{\theta}_j + \gamma a_j \left( G_j \right) \left( \tilde{z}_j + \tilde{\zeta}_j \right),
\]

\[
\tilde{R}_M = \sum_{j=1}^{N+K} (\tilde{z}_j + \tilde{\zeta}_j) \tilde{R}_j = \sum_{j=1}^{N+K} \left[ (\tilde{z}_j + \tilde{\zeta}_j) \left( \tilde{\theta}_j + \gamma a_j \left( G_j \right) \left( \tilde{z}_j + \tilde{\zeta}_j \right) \right) \right].
\]

It follows that:

\[
\text{Var} \left( \tilde{R}_M \right) = \sum_{j=1}^{N+K} \text{Var} \left[ \left( \tilde{z}_j + \tilde{\zeta}_j \right) \left( \tilde{\theta}_j + \gamma a_j \left( G_j \right) \left( \tilde{z}_j + \tilde{\zeta}_j \right) \right) \right]
\]

\[
= \sum_{j=1}^{N+K} \text{Var} \left[ \tilde{z}_j \tilde{\theta}_j + 2 \tilde{z}_j \gamma a_j \left( G_j \right) \tilde{z}_j + \tilde{\zeta}_j \tilde{\theta}_j + \gamma a_j \left( G_j \right) \tilde{\zeta}_j^2 \right]
\]

\[
= \sum_{j=1}^{N+K} \left[ E \left( \left( \tilde{z}_j \tilde{\theta}_j + 2 \tilde{z}_j \gamma a_j \left( G_j \right) \tilde{z}_j + \tilde{\zeta}_j \tilde{\theta}_j + \gamma a_j \left( G_j \right) \tilde{\zeta}_j^2 \right)^2 \right) \right. 
\]

\[- \left. \left[ E \left( \tilde{z}_j \tilde{\theta}_j + 2 \tilde{z}_j \gamma a_j \left( G_j \right) \tilde{z}_j + \tilde{\zeta}_j \tilde{\theta}_j + \gamma a_j \left( G_j \right) \tilde{\zeta}_j^2 \right) \right] \right]^2 \right]
\]

\[
= \sum_{j=1}^{N+K} \left[ \left( v_{0j} + 4 \gamma^2 a_j \left( G_j \right)^2 v_{0j} \right) \tilde{z}_j^2 + v_{0j} v_{0j} + 3 \gamma^2 a_j \left( G_j \right)^2 \tilde{z}_j^2 - \gamma^2 a_j \left( G_j \right)^2 \tilde{z}_j^2 + 2 \gamma^2 a_j \left( G_j \right)^2 \tilde{z}_j^2 \right]
\]

\[
= \sum_{j=1}^{N+K} \left[ \left( v_{0j} + 4 \gamma^2 a_j \left( G_j \right)^2 v_{0j} \right) \tilde{z}_j^2 + v_{0j} v_{0j} + 3 \gamma^2 a_j \left( G_j \right)^2 \tilde{z}_j^2 - \gamma^2 a_j \left( G_j \right)^2 \tilde{z}_j^2 \right].
\]

Note that in the third equality, all cross-products of items in the first bracket have mean zero. In the fourth equality, \( E(\tilde{z}_j^2) = 3 \nu_{0j}^2 \).

Computation of the covariance between the \( j \)th basic security’s and market portfolio’s returns from the definition in Section 2.3

From the expressions of \( \tilde{R}_j \) and \( \tilde{R}_M \):

\[
\text{Cov} \left( \tilde{R}_j, \tilde{R}_M \right) = \text{Cov} \left( \tilde{R}_j, \sum_{j=1}^{N+K} (\tilde{z}_j + \tilde{\zeta}_j) \tilde{R}_j \right)
\]

\[
= \text{Cov} \left( \tilde{R}_j, (\tilde{z}_j + \tilde{\zeta}_j) \tilde{R}_j \right)
\]

\[
= \text{Cov} \left( \tilde{R}_j, \tilde{z}_j \tilde{R}_j \right) + \text{Cov} \left( \tilde{R}_j, \tilde{\zeta}_j \tilde{R}_j \right),
\]

where the second equality obtains because basic securities have independent random supplies and payoffs:

\[
\text{Cov} \left( \tilde{R}_j, \tilde{z}_j \tilde{R}_j \right) = \tilde{z}_j \text{Var} \left( \tilde{R}_j \right) = v_{0j} \tilde{z}_j^2 + \gamma^2 a_j \left( G_j \right)^2 v_{0j} \tilde{z}_j.
\]
\[
\text{Cov} \left( \tilde{R}_j, \tilde{z}_j \tilde{R}_j \right) = E \left( \tilde{z}_j \tilde{R}_j^2 \right) - E \left( \tilde{R}_j \right) E \left( \tilde{z}_j \tilde{R}_j \right)
\]

\[
= E \left[ \tilde{z}_j (\tilde{\theta}_j + \gamma a_j(G_j) (\tilde{\xi}_j + \tilde{z}_j)) \right]^2
- E \left[ \tilde{\theta}_j + \gamma a_j(G_j) (\tilde{\xi}_j + \tilde{z}_j) \right] E \left[ \tilde{z}_j (\tilde{\theta}_j + \gamma a_j(G_j) (\tilde{\xi}_j + \tilde{z}_j)) \right]
\]

\[
= 2\gamma^2 a_j(G_j)^2 v_z \tilde{z}_j - \gamma^2 a_j(G_j)^2 v_z \tilde{\xi}_j
\]

\[
= \gamma^2 a_j(G_j)^2 v_z \tilde{\xi}_j.
\]

Therefore, \( \text{Cov}(\tilde{R}_j, \tilde{R}_M) = v_\theta \tilde{\xi}_j + 2\gamma^2 a_j(G_j)^2 v_z \tilde{z}_j. \)

**Appendix 3**

Our analysis in Section 3 focused on the basic securities for tractability. We now show that the main results of that section (P6 and P7) carry over to the original securities.

**Subsection 3.1**

Consider the all-non-G economy. We can use Equation (2) to reconstruct the original securities using the basic securities. Specifically, consider the \( j \)th (\( j \in 1, \ldots, N \)) non-factor original security as a portfolio of \( \tilde{V}_j \) units of the risk-free asset, \( \tilde{\beta}_j \) \( (k = 1, \ldots, K) \) units of the \( N + k \)th basic security (the \( k \)th factor, \( \tilde{\xi}_j \)), and one unit of the \( j \)th basic security (the \( j \)th residual, \( \tilde{z}_j \)).

We can use an analysis similar to that in the proof of Lemma 3 to show that the price of the \( j \)th non-factor original security is given by:

\[
P^{A_NG}(V_j) = \tilde{V}_j + \sum_{k=1}^{K} \tilde{\beta}_j P_{N+k}^{A_NG} + P_j^{A_NG},
\]

where \( P_{N+k}^{A_NG} \) is the price of the \( N + k \)th basic security (the \( k \)th factor), and \( P_j^{A_NG} \) is the price of the \( j \)th basic security (the \( j \)th residual) as given in Equation (20). The return of the original security can be expressed as:

\[
\tilde{R}_j^{A_NG} = \sum_{k=1}^{K} \tilde{\beta}_j \left[ \tilde{\theta}_{N+k} + \gamma \tilde{v}_{N+k} (\tilde{\xi}_{N+k} + \tilde{z}_{N+k}) \right] + \left[ \tilde{\theta}_j + \gamma \tilde{v}_0 (\tilde{\xi}_j + \tilde{z}_j) \right].
\]

(A18)

The expected return and volatility of the original security are given by:

\[
E \left( \tilde{R}_j^{A_NG} \right) = \sum_{k=1}^{K} \tilde{\beta}_j \left[ \gamma \tilde{v}_{N+k} (\tilde{\xi}_{N+k} + \tilde{z}_{N+k}) \right] + \left[ \gamma \tilde{v}_0 \tilde{z}_j \right],
\]

(A19)

\[
\text{Var} \left( \tilde{R}_j^{A_NG} \right) = \sum_{k=1}^{K} \tilde{\beta}_j^2 \left[ \gamma \tilde{v}_{N+k} + (\gamma \tilde{v}_{N+k})^2 \tilde{v}_{N+k} \right] + \left[ \gamma \tilde{v}_0 + (\gamma \tilde{v}_0)^2 \tilde{v}_j \right].
\]

One can estimate the market \( \beta \) by regressing \( \tilde{R}_j^{A_NG} \) against the return on the market portfolio. Note that the market portfolio of the original securities is just a reshuffle of the basic securities, and is therefore identical to the market portfolio of the basic securities. It follows from Equations (7) (we just need to replace \( a_j(G_j) \) by \( \tilde{v}_\theta \)) and (A18) that the market \( \beta \) of the original security is given by:

\[
\beta^{A_NG}_M = \frac{\sum_{k=1}^{K} \tilde{\beta}_j \left[ \gamma \tilde{v}_{N+k} + 2\gamma^2 \tilde{v}_{N+k} \tilde{v}_{N+k} \right] \tilde{\xi}_{N+k} + \left[ \gamma \tilde{v}_0 + 2\gamma^2 \tilde{v}_{N+k} \tilde{v}_j \right] \tilde{z}_j}{\text{Var} \left( \tilde{R}_M^{A_NG} \right)}.
\]
Write the expected return of the original security (see Equation (A19)) as:

\[
E(\tilde{R}_j^{A,NG}) = \tilde{\lambda}_j^{A,NG} \beta_j^{A,NG},
\]

where:

\[
\tilde{\lambda}_j^{A,NG} = \frac{\sum_{k=1}^{K} \beta_j[k(\gamma \nu_{k+j} + \tilde{\zeta}_{N+k}) + \gamma \nu_{k+j} \tilde{\xi}_j]}{\sum_{k=1}^{K} \beta_j[kv_{k+N+j} + 2\gamma^2v_{k+N+j}v_{N+k} + \nu_{k+j}v_{k+j}]} \text{Var}(\tilde{R}_M^{A,NG}).
\]

Consider the all-$G$ economy studied in Section 3.1. We can use the same analysis as above to show that the return of the $j$th non-factor original security can be expressed as:

\[
R_j^{A,G} = \sum_{k=1}^{K} \beta_j[k(\gamma (v_{k+N+j} - G_{N+k}/\gamma^2) \tilde{\zeta}_{N+k}) + (\gamma (v_{k+N+j} - G_{N+k}/\gamma^2))] + [\gamma (v_{k+N+j} - G_{N+k}/\gamma^2)](\tilde{\xi}_j + \tilde{\xi}_j).
\]

The expected return and volatility of the original security are given by:

\[
E(\tilde{R}_j^{A,G}) = \frac{\sum_{k=1}^{K} \beta_j[k(\gamma (v_{k+N+j} - G_{N+k}/\gamma^2) \tilde{\zeta}_{N+k}) + (\gamma (v_{k+N+j} - G_{N+k}/\gamma^2))] + [\gamma (v_{k+N+j} - G_{N+k}/\gamma^2)](\tilde{\xi}_j + \tilde{\xi}_j)}{\text{Var}(\tilde{R}_M^{A,G})}
\]

The market $\beta$ of the original security is given by:

\[
\beta_j^{A,G} = \frac{\sum_{k=1}^{K} \beta_j[kv_{k+N+j} + 2\gamma^2(v_{k+N+j} - G_{N+k}/\gamma^2)v_{N+k} + \nu_{k+N+j} + 2\gamma^2(v_{k+N+j} - G_{N+k}/\gamma^2)v_{k+j}]}{\text{Var}(\tilde{R}_M^{A,G})}
\]

Write the expected return of the original security as:

\[
E(\tilde{R}_j^{A,G}) = \tilde{\lambda}_j^{A,G} \beta_j^{A,G},
\]

where:

\[
\tilde{\lambda}_j^{A,G} = \frac{\sum_{k=1}^{K} \beta_j[k(\gamma (v_{k+N+j} - G_{N+k}/\gamma^2) \tilde{\zeta}_{N+k}) + (\gamma (v_{k+N+j} - G_{N+k}/\gamma^2))] \text{Var}(\tilde{R}_M^{A,G})}{\sum_{k=1}^{K} \beta_j[kv_{k+N+j} + 2\gamma^2(v_{k+N+j} - G_{N+k}/\gamma^2)v_{N+k} + \nu_{k+N+j} + 2\gamma^2(v_{k+N+j} - G_{N+k}/\gamma^2)v_{k+j}]} \tilde{\xi}_j
\]

It is straightforward to show from the above analysis that for the original security, $E(\tilde{R}_j^{A,NG}) > E(\tilde{R}_j^{A,G})$ and $\text{Var}(\tilde{R}_j^{A,NG}) > \text{Var}(\tilde{R}_j^{A,G})$ (this is consistent with P6 (i)).

Note that our result in P6 (ii) on the market portfolio continues to hold for the original securities. The reason is that the market portfolio of the original securities is just a reshuffle of the basic securities, and is therefore identical to the market portfolio of the basic securities.

To facilitate a simple comparison between $\tilde{\lambda}_j^{A,NG}$ and $\tilde{\lambda}_j^{A,G}$, let $\forall k G_{N+k} = 0$ (i.e. $G$ traders receive no extra utility from trading the factors). Then, we can use a similar analysis as in the proof of P6 (iii) to show that $\tilde{\lambda}_j^{A,NG} > \tilde{\lambda}_j^{A,G}$ if $\forall k v_{k+N+j}v_{N+k} \geq (\nu_{k+N+j} - (G_{N+k}/2\gamma^2))v_{k+j}$.

**Subsection 3.2**

In this hybrid case, $G$ traders are present in the trading of the first $N_1$ non-factor original securities and all the $N+k$th ($k = 1, \ldots, K$) original securities (this is equivalent to trading the first $N_1$ residuals and all the $K$ factors). They are not present in the remaining $N_N$ non-factor original securities (or the remaining $N-N_1$ residuals).
We can apply an analysis similar to the above to show that the return of the $j$th non-factor original security with $G$ traders’ presence can be expressed as:

\[
\hat{R}_j = \sum_{k=1}^{K} \beta_{jk} \left[ \hat{\theta}_{N+k} + \gamma a_{N+k}(G_{N+k})(\hat{x}_{N+k} + \hat{z}_{N+k}) \right] + \left[ \hat{\theta}_j + \gamma (G_j)(\hat{x}_j + \hat{z}_j) \right].
\] (A20)

The expected return of the original security is given by:

\[
E\left( \hat{R}_j \right) = \sum_{k=1}^{K} \beta_{jk} [\gamma a_{N+k}(G_{N+k})\hat{x}_{N+k}] + [\gamma a_j (G_j)\hat{x}_j].
\] (A21)

It follows from Equations (7) and (A20) that the market \( \beta \) of the original security is:

\[
\beta_{jM} = \frac{\sum_{k=1}^{K} \beta_{jk} \left[ v_{N+k} + 2\gamma^2 a_{N+k}(G_{N+k})^2 v_{N+k} \right] \hat{x}_{N+k} + \left[ v_{0j} + 2\gamma^2 a_j (G_j)^2 v_{0j} \right] \hat{x}_j}{\text{Var}\left( \hat{R}_M \right)}.
\]

Write the expected return of the original security (see Equation (A21)) as:

\[
E\left( \hat{R}_j \right) = \lambda_j \beta_{jM},
\]

where:

\[
\lambda_j = \frac{\sum_{k=1}^{K} \beta_{jk} [\gamma a_{N+k}(G_{N+k})\hat{x}_{N+k}] + [\gamma a_j (G_j)\hat{x}_j]}{\sum_{k=1}^{K} \beta_{jk} \left[ v_{N+k} + 2\gamma^2 a_{N+k}(G_{N+k})^2 v_{N+k} \right] \hat{x}_{N+k} + \left[ v_{0j} + 2\gamma^2 a_j (G_j)^2 v_{0j} \right] \hat{x}_j} \text{Var}\left( \hat{R}_M \right).
\]

The return of the $j$th non-factor original security without the presence of $G$ traders is given by:

\[
\hat{R}_j^o = \sum_{k=1}^{K} \beta_{jk} \left[ \hat{\theta}_{N+k} + \gamma a_{N+k}(G_{N+k})(\hat{x}_{N+k} + \hat{z}_{N+k}) \right] + \left[ \hat{\theta}_j + \gamma (v_{0j} / \rho)(\hat{x}_j + \hat{z}_j) \right].
\] (A22)

The expected return of the original security is given by:

\[
E\left( \hat{R}_j^o \right) = \sum_{k=1}^{K} \beta_{jk} [\gamma a_{N+k}(G_{N+k})\hat{x}_{N+k} + \gamma (v_{0j} / \rho)\hat{x}_j].
\] (A23)

It follows from Equations (7) and (A22) that the market \( \beta \) of the original security is:

\[
\beta_{jM}^o = \frac{\sum_{k=1}^{K} \beta_{jk} \left[ v_{N+k} + 2\gamma^2 a_{N+k}(G_{N+k})^2 v_{N+k} \right] \hat{x}_{N+k} + \left[ v_{0j} + 2\gamma^2 (v_{0j} / \rho)^2 v_{0j} \right] \hat{x}_j}{\text{Var}\left( \hat{R}_M \right)}.
\]

Write the expected return of the original security (see Equation (A23)) as:

\[
E\left( \hat{R}_j^o \right) = \lambda_j^o \beta_{jM}^o.
\]
where:

$$x_j^o = \frac{\sum_{k=1}^{K} B_{jk} [\gamma a_{N+k}(G_{N+k})\tilde{\xi}_{N+k}] + [\gamma (v_{0}/\rho)\tilde{\xi}_j]}{\sum_{k=1}^{K} B_{jk} [v_{0N+k} + 2\gamma^2 a_{N+k}(G_{N+k})^2 v_{2N+k}] \tilde{\xi}_{N+k} + [v_{0} + 2\gamma^2 (v_{0}/\rho)^2 v_{2j}] \tilde{\xi}_j} \Var(\tilde{R}_M).$$

We can use a similar analysis as in the proof of P7 to show that for the two original securities, $j$ (with the presence of $G$ traders) and $j'$ (without the presence of $G$ traders), $\lambda_j > \lambda_j^o$, $\forall k a_{N+k}(G_{N+k})v_{2N+k} > ((v_{0}/\rho + (G_j))/2)v_{2j}$.

**Corresponding author**

Avanidhar Subrahmanyam can be contacted at: subra@anderson.ucla.edu