# Market Size Matters:

# A Model of Excess Volatility in Large Markets*

Kei Kawakami†

March 9th, 2015

**Abstract**

We present a model of excess volatility based on speculation and equilibrium multiplicity. Each trader has two distinct motives to trade: (i) speculation based on noisy signals, and (ii) hedging against endowment shocks. The key to equilibrium multiplicity is the self-fulfilling nature of information aggregation: if individuals trade relatively more on the basis of speculation rather than hedging, then prices reveal more information on payoff risk which in turn justifies less need for hedging. We first show that multiplicity arises only in large markets where aggregate shocks in prices are sufficiently more important than idiosyncratic shocks. We then show that (i) in a given equilibrium, excess volatility increases with payoff volatility, (ii) comparing across different equilibria, excess volatility is negatively associated with liquidity, trading volume, and social welfare. We also show that an increase in market size either creates high-volatility equilibria or eliminates low-volatility equilibria. Among other things, the model predicts that given two assets identical in their fundamental properties, the one that attracts more traders overtime is more likely to experience a jump in excess volatility.

**Keywords**: Asymmetric information, Excess volatility, Multiple equilibria, Price impact, Volume, Welfare.

*The author acknowledges detailed comments from Chris Edmond. All errors are mine.

†Department of Economics, University of Melbourne, e-mail: keik@unimelb.edu.au. The Appendix is available at my website: http://sites.google.com/site/econkeikawakami/research

# 1 Introduction

Excess volatility – price movements not easily explained by changes in fundamentals – seems prevalent in modern stock markets (West 1988, Shiller 2003, 2014). Moreover, extreme episodes of excess volatility, such as bubbles and crises, are observed even in markets with a very large number of traders. Intuitively, one might think that excess volatility should be less likely in large markets. For example, in large markets idiosyncratic shocks cancel out and each trader's influence on prices (i.e., price impact) will be negligible.[1] But on the other hand, it seems natural that coordination effects will be larger when there are many traders. Moreover coordination is more difficult to sustain precisely when many traders interact anonymously. In large markets where coordination becomes *powerful but fragile*, equilibrium multiplicity might be relevant for excess volatility. The aim of this paper is to understand the role of market size for excess volatility in this context.

We present a model based on the trading game that Kyle (1989) and others used to study asset pricing when risk averse and strategic traders have diverse private information about the payoff value of the asset. Unlike Kyle (1989), however, each trader has two distinct motives to trade: (i) speculation based on privately observed noisy signals, and (ii) hedging needs for his endowment position. Signal noise can be interpreted as non-fundamental shocks to beliefs about asset *quality*, while endowment positions represent *quantity* shocks. Both shocks can be correlated among traders. We define excess volatility as the variance of prices conditional on the payoff value and aggregate endowment of the asset. Excess volatility exists because speculation causes prices to fluctuate due to belief shocks unrelated to fundamentals. The key to equilibrium multiplicity is the self-fulfilling nature of information aggregation: coordinating on more speculation and less hedging makes prices more revealing about payoff risk, which in turn justifies less need for hedging, and vice-versa.

In the model, cross-sectional correlation in belief shocks and in endowment shocks act as two aggregate shocks: a shock to the aggregate belief (i.e., pure noise in the average belief),

[1] Gabaix et al (2006) proposes a model of excess volatility caused by price impact of large traders.

and a shock to the aggregate endowment. Importantly, both shocks affect prices but cannot be separately disentangled, and rational traders estimate them by using their private information. Thus, each piece of private information has a *secondary role* in estimating different aggregate shocks in prices, *in addition to its primary role.* Importantly, the secondary role works as a discounting factor of the primary role. To see this, consider the dual roles for a signal. First, the signal is used to estimate the payoff value of the asset (a primary role). Second, because prices aggregate individual signals, a high realization of prices can be either due to the high payoff value of the asset or due to a high realization of the common shock in the signal. When traders estimate the impact of the common shock in prices by using their signal (a secondary role), the overall impact of the signals on traders' orders is reduced. Similarly, the individual endowment also has dual roles. A high realization of the endowment makes traders want to hedge the asset's payoff risk (a primary role). In the mean time, a high realization of the endowment implies that prices are low in part due to a high realization of the common shock in endowments. Again, when traders estimate the impact of the common shock by using their individual endowment (a secondary role), the overall impact of endowment on traders' orders is reduced. In circumstances where the common shocks in prices are important, the secondary roles of signals and endowment become more important relative to their primary roles.

We first establish conditions for multiplicity (**Proposition 1**). In particular, we show that multiplicity occurs (a) only if the cross-sectional correlation of endowment shocks is sufficiently high relative to the cross-sectional correlation of belief shocks, (b) only in markets that are sufficiently large, and (c) given (a) and (b), multiplicity occurs if model parameters do not imply that one motive is too dominant relative to the other motive.

To see why condition (a) is necessary for multiplicity, first note that the condition states that the aggregate endowment shock is more important than the aggregate belief shock in prices. Suppose that a given trader conjectures that other traders increase their level of speculation relative to hedging, which implies that he can potentially extract more information from prices. This makes the secondary roles of his signal and endowment more important, but relatively more so for his endowment under condition (a). Accordingly, this trader reduces hedging relative to speculation. This implies that a conjecture that other traders speculate more (relative to hedging) makes a given trader want to do the same, i.e., traders' *motives are complements*. By contrast, when endowment shocks are less correlated than belief shocks, the same conjecture makes a given trader want to do the opposite, i.e., motives are substitutes.[2]

Condition (b) is necessary for multiplicity, because market size matters for the degree of complementarity/substitutability in motives. Because the variances of *averaged idiosyncratic shocks* in prices decrease with market size, the relative importance of the common shocks in prices increases with market size. Therefore, given that complementarity or substitutability in trading motives exits, it becomes even stronger if the market is large. Because sufficiently strong complementarity in trading motives is necessary for multiplicity, it arises only in large markets. Finally, condition (c) indicates an "ambiguous" environment that is not too conductive to one motive. This ensures that the complementarity in motives has a *decisive* impact on traders' choice of strategies.[3]

To isolate the role of multiplicity, we first study excess volatility in a given equilibrium. We show that while prices underreact to the innovation in the payoff value of the asset, excess volatility increases with payoff volatility (**Proposition 2**). This is because excess volatility in our model comes from noise in signals, i.e., non-fundamental belief shocks. On the one hand, noise in signals softens the reaction of prices to the change in the payoff value, because part of it is attributed to noise. On the other hand, when the payoff value is very volatile, traders put more weight on signals, thereby increasing the importance of belief shocks for

[2] Our discussion here is not based on a formal ordering of strategies. In our model, a strategy is a mapping from private information to a demand function, which can be viewed as a random demand function. We do not analyze the ordering of general random functions. However, we intoroduce the measure of the balance of two motives in the class of linear demand functions. According to this ordering, our notion of substitutes/complements in *motives* coincides with the standard notion of strategic substitutes/complements.

[3] A similar condition which ensures a non-trivial situation arises in many coordination games. For example, in a standard coordination game used in the global game literature, the strength of a regime must be in an intermediate region to make coordination a relevant issue.

prices. Thus, our model can explain two seemingly unrelated phenomena — excess volatility and the underreaction of prices to fundamentals — in a simple learning framework.

We then go on to characterize key properties of the equilibria. Importantly, we show that equilibria can be ranked by the level of excess volatility. Across multiple equilibria (i.e., for fixed model parameters including market size), the level of excess volatility is negatively associated with liquidity[4], trade volume, and aggregate welfare (**Proposition 3**). This ranking is intuitive. In an equilibrium where traders coordinate on more speculation relative to hedging, belief shocks matter more for prices. At the same time, because now there is more information to be learned from prices, the importance of estimating aggregate shocks in prices increases. As a result, traders discount the primary role of private information by more, leading to smaller liquidity and volume.

Like any model with multiple equilibria, the model proposed here suggests that equilibrium outcomes, including excess volatility, can be sensitive to a small change in the underlying environment (e.g., market size, preferences of traders, and information structure). We go one-step further by studying which type of changes can support a desirable choice of equilibrium by either eliminating bad equilibria or creating good equilibria. We show that an increase in market size cannot achieve this (**Proposition 4**). More precisely, if an increase in market size creates or eliminates an equilibrium, it must also be the case that: 1) bad equilibria emerge, or 2) good equilibria disappear. In this sense, the model indicates that an increase in market size may induce an equilibrium selection that is *detrimental* to traders.

Section 2 describes the model. We first establish the existence of equilibrium in large markets. Section 3 studies how and when multiplicity arises. Section 4 characterizes the key properties of equilibria and discusses empirical implications and related literature. Section 5 concludes. The appendix contains all proofs.[5]

[4] We measure liquidity by the inverse of *price impact*, i.e., the extent to which each trader can affect prices.
[5] The Appendix is available at my website: http://sites.google.com/site/econkeikawakami/research.

# 2 Model

**Endowment and information.** There are $n+1$ traders indexed by $i \in \{1, ..., n+1\}$ who trade endowments that have an unknown payoff $v$ per unit. The payoff $v$ has a distribution $N(0, \tau_v^{-1})$. Before trading, each trader $i$ receives (i) the endowment $e_i$, and (ii) a private signal $s_i$ about $v$. Both are private information. There is no additional supply beyond the sum of endowments $\sum_{i=1}^{n+1} e_i$. The endowment has a common component $x_0$ and an idiosyncratic component $x_i$:

$$e_i = \sqrt{1-u}\ x_0 + \sqrt{u}\ x_i, \qquad u \in [0,1], \tag{1}$$

where $x_0$ and $x_i$ are independently normally distributed with mean zero and variance $\tau_x^{-1}$. Trader $i$ knows the realized $e_i$ but does not observe the components $x_0$ and $x_i$ separately. The parameter $u$ in (1) determines the relative importance of the two components. We focus on $u > 0$ so that hedging opportunities exist.[6]

Each trader observes a noisy signal of the payoff $v$ of the form $s_i = v + \varepsilon_i$. We allow for cross-sectional correlation in the noise $\varepsilon_i$:

$$\varepsilon_i = \sqrt{1-w}\ \epsilon_0 + \sqrt{w}\ \epsilon_i, \qquad w \in [0,1], \tag{2}$$

where $\epsilon_0$ and $\epsilon_i$ are independently normally distributed with mean zero and variance $\tau_\varepsilon^{-1}$. The parameter $w$ in (2) measures *how differentially informed* traders are. If $w$ is zero, the signals are identical and the information is symmetric. As $w$ increases, information is more heterogeneous and there are more gains from information aggregation. We emphasize here that signal noise $\varepsilon_i$, (2), is the only exogenous source of excess volatility in our model. While variance of $\varepsilon_i$ is fixed, the extent to which it affects prices depends on the way traders use signals in their trading, which is endogenously determined.

[6] If $u$ is zero, there is no diversifiable risk and no trade can happen.

For any pair of traders $i \neq j$, the *endowment correlation* is $Corr[e_i, e_j] = 1 - u$ and the *belief shock correlation* is $Corr[\varepsilon_i, \varepsilon_j] = 1 - w$. If $u = w$, the correlations are identical.[7] The specification (1) and (2) ensures that ex ante variances of $e_i$ and $\varepsilon_i$ do not depend on $(u, w)$.

**Preferences.** Each trader has an exponential utility function $U(\pi_i) = -\exp(-\rho\pi_i)$ with a risk-aversion coefficient $\rho > 0$, where $\pi_i$ is trader $i$'s profit. Let $p$ denote a market-clearing price. Profits are the the payoff from the post-trade position $q_i + e_i$ less the payment for trading $pq_i$, and hence $\pi_i = v(q_i + e_i) - pq_i$.[8] After observing the private information $H_i = (e_i, s_i)$, each trader chooses her order $q_i(p; H_i)$.

**Equilibrium.** We characterize a *linear Bayesian Nash equilibrium*, where each trader submits a linear demand function

$$q_i(p; e_i, s_i) = \beta_s s_i - \beta_e e_i - \beta_p p \tag{3}$$

for some constants $(\beta_s, \beta_e, \beta_p)$ to be determined in equilibrium. A market-clearing price $p^*$ satisfies

$$\sum_{i=1}^{n+1} q_i(p^*; H_i) = 0. \tag{4}$$

Because an order can be explicitly conditioned on prices $p$, traders can internalize the informational content of $p^*$, i.e., they are Bayesian. Moreover, each trader internalizes how his order affects $p^*$ through (4), i.e., they are Nash players rather than price-takers.[9] Trader $i$ obtains the quantity determined by his order evaluated at the market-clearing price, $q_i^* = q_i(p^*; H_i)$, for which he pays $p^* q_i^*$. Following the literature, we characterize this equilibrium by a guess-and-verify method.

[7] If $(u, w) = (1, 1)$, both $e_i$ and $\varepsilon_i$ are i.i.d. This case was studied by Diamond and Verecchia (1981) assuming price-taking traders.

[8] $q_i$ can be negative, in which case trader $i$ receives $-pq_i$.

[9] This solution concept was first proposed by Kyle (1989). Vives (2009) and Rostek and Weretka (2012) are recent works based on this solution concept. Kyle (1989) in fact allows demand correspondence to be used, but specifies a market-clearing rule such that functions are submitted in equilibrium. We abstract from the technical issues by directly assuming that traders submit demand functions, and that no trade occurs if no price (or multiple prices) can clear the market.

We expect $(\beta_s, \beta_e, \beta_p)$ to be positive in equilibrium, and call $\beta_s$ the *speculation intensity* and call $\beta_e$ the *hedging intensity*. We measure the balance of the two trading motives by the ratio of $\beta_e$ to $\beta_s$. It turns out that normalizing this ratio by multiplying $\frac{\tau_\varepsilon}{\rho}$ simplifies the exposition. Therefore, we define the balance of the two trading motives by

$$k \equiv \frac{\tau_\varepsilon}{\rho} \frac{\beta_e}{\beta_s}.$$

Note that $k$ is an endogenous variable. If equilibrium orders (3) exhibit a high level of hedging intensity $\beta_e$ relative to speculation intensity $\beta_s$, $k \equiv \frac{\tau_\varepsilon}{\rho} \frac{\beta_e}{\beta_s}$ becomes large, in which case we say that equilibrium involves more hedging relative to speculation. The following notation turns out to be helpful to discuss our key results:

$$\alpha_\varepsilon \equiv \frac{\rho^2}{\tau_\varepsilon \tau_x}, \qquad N_u \equiv 1 + (1-u)\, n, \qquad N_w \equiv 1 + (1-w)\, n. \tag{5}$$

These are exogenous parameters. First, $\alpha_\varepsilon$ measures whether the environment is relatively more conductive to speculation or more conductive to hedging. The larger is $\alpha_\varepsilon$, the more the environment is favorable to hedging relative to speculation. Second, $N_u - 1$ and $N_w - 1$ measure the size of aggregate shocks from an individual trader's perspective. Importantly, market size affects an equilibrium information structure through $\frac{N_w}{N_u}$.

Given the assumptions on preferences and random variables, it can be shown that trader $i$ holds a belief in equilibrium that $v$ has a distribution $N\left(E_i[v], \frac{1}{\tau}\right)$, where $E_i[\cdot]$ is the conditional expectation operator $E[\cdot|e_i, s_i, p]$ and $\tau$ is the precision of beliefs.[10] Both the mean $E_i[v]$ and the precision $\tau$ are to be determined in equilibrium. Trader $i$'s problem is

$$\max_{q_i} \left\{ E_i[v]\,(q_i + e_i) - \frac{\rho}{2\tau}\,(q_i + e_i)^2 - p q_i \right\} \tag{6}$$

$$\text{subject to} \qquad p = p_i + \lambda q_i, \tag{7}$$

[10] $\tau$ is independent of $i$ because of linear-normal information structure and ex ante symmetry of traders.

where we define $p_i \equiv \frac{1}{n\beta_p}\left(\beta_s \sum_{j\neq i} s_j - \beta_e \sum_{j\neq i} e_j\right)$ and denote the *price impact* $\lambda \equiv \frac{1}{n\beta_p}$. To see where the constraint (7) comes from, rewrite the market-clearing condition (4) using the conjectured order (3):

$$q_i^* + \sum_{j\neq i} q_j\left(p^*; H_j\right) = 0 \Leftrightarrow q_i^* + \left(\beta_s \sum_{j\neq i} s_j - \beta_e \sum_{j\neq i} e_j\right) = n\beta_p p^*.$$

By dividing both sides by $n\beta_p$, (7) is obtained. Without the constraint (7), $p$ in (6) can be any value independent of $q_i$, i.e., it is just a parameter from individual trader's perspective. The constraint (7) means that trader $i$ is aware that $p$ in (6) must be $p^*$ that satisfies the equilibrium condition (4).[11] Solving this problem, the optimal order of trader $i$ is

$$q_i\left(p\right) = \frac{E_i[v] - p - \frac{\rho}{\tau}e_i}{\lambda + \frac{\rho}{\tau}}. \tag{8}$$

To completely solve the problem, we need to characterize the beliefs, namely the mean $E_i[v]$ and the precision $\tau$. First, beliefs depend on prices upon which the order is contingent, because each realization of prices has a different implication for the distribution of $v$. Second, prices depend on strategies. Therefore, beliefs in (8) depend on the conjecture (3). It turns out that the optimal order (8) is indeed linear in $(s_i, e_i, p)$:

$$q_i\left(p; e_i, s_i\right) = \widetilde{\beta}_s s_i - \widetilde{\beta}_e e_i - \widetilde{\beta}_p p. \tag{9}$$

Taken together the conjectured order (3) and the best response order (9) imply a fixed point problem in the coefficients $\left(\beta_s, \beta_e, \beta_p\right)$. Because this is a standard procedure in the literature, we gather details in the Appendix. We focus on the case where a strategic foundation to large markets can be provided.[12]

[11] Price-taking equilibrium is obtained by ignoring the constraint (7) and treating $p$ as a parameter. While this assumption is not consistent with finite $n$, we can show that our main results (except those for price impact) still hold under this alternative solution concept. See the Appendix for more details.

[12] It is well known that a linear equilibrium may fail to exist for small $n$ when traders act strategically. If we assume price-taking behavior, equilibrium exists for all $n \geq 1$. See the Appendix for more details.

**Lemma 1 (equilibrium existence for large markets)**

*If $w < 1$, then an equilibrium exists for sufficiently large $n$.*

The condition $w < 1$ is sufficient for the existence of a linear equilibrium in large markets, but not necessary.[13] While we obtain some interesting results for the case with $w = 1$ (i.e., no correlation in belief shocks), these results are not robust to a small change in $w$.[14] Therefore, we focus on the case with $w < 1$. In the next section, we study the endogenous determination of beliefs $\left(E_i[v], \frac{1}{\tau}\right)$, and how and when they generate multiplicity.

# 3 Equilibrium multiplicity

To understand how equilibrium multiplicity arises in our model, we devote the next subsection to the analysis of complementarities in trading motives that endogenously arises in our model. This serves as a stepping stone to appreciating the precise conditions for multiplicity, which are presented in the following subsection.

## 3.1 Complementarity in trading motives

We proceed in four steps. First, we introduce a particular measure of information aggregation, $\varphi \in (0, 1)$. The way in which $\varphi$ is determined plays an important role for equilibrium multiplicity. Second, to highlight why the endogenous determination of $\varphi$ is crucial, we first suppose that traders conjecture some fixed value $\varphi$ independent of the conjectured strategies (3). We call this a *naive conjecture* of $\varphi$, because it is typically inconsistent with optimal strategies. Third, we analyze how $\varphi$ is determined when traders form a *rational conjecture* of $\varphi$. Finally, we study how market size $n$ affects the nature of the rational conjecture.

[13] If $u = w = 1$, an equilibrium exists for sufficiently large $n$ given the additional condition $\alpha_\varepsilon > 1$.

[14] See also Manzano and Vives (2011). Ganguli and Yang (2009) studied a case with $w = 1 > u$ with a continuum of traders, and found that two linear equilibria exist if $\alpha_\varepsilon$ is above a certain threshold, while there is no linear equilibrium if $\alpha_\varepsilon$ is below it. In the Appendix, we prove the following results. First, if $w = 1 > u$ and $\alpha_\varepsilon \geq \frac{4}{u}$, then there are two linear equilibria for large $n$. Second, if $w = 1 > u$ and $\alpha_\varepsilon < \frac{4}{u}$, equilibrium may exist for a finite $n$, but even if it exists, it disappears as $n$ increases to a *finite* threshold. This is the only case where price impact *increases* with market size.

To derive the right measure of information aggregation, begin with the conjectured order (3) and use the market-clearing condition (4) to see that equilibrium prices take the form:

$$p = \frac{\beta_s}{\beta_p} v - \frac{\beta_e}{\beta_p} \left( \sqrt{1-u} x_0 + \sqrt{u} \overline{x} \right) + \frac{\beta_s}{\beta_p} \sqrt{1-w} \epsilon_0 + \sqrt{w} \overline{\epsilon}, \tag{10}$$

where $\overline{x}$ is the average of $x_i$ and $\overline{\epsilon}$ is the average of $\epsilon_i$. This formula for prices has two features. First, as long as $w > 0$, it provides information about $v$ that cannot be learned from individual's own signal $s_i$. Second, it is correlated with individual variables $(e_i, s_i)$ both through the common components $(x_0, \epsilon_0)$ and the idiosyncratic components $(x_i, \epsilon_i)$. Importantly, the *relative* correlations of $(s_i, e_i)$ with prices (10) depend on $\left( \frac{\beta_e}{\beta_s}, u, w, n \right)$. With this in mind, we now study information aggregation by prices.

**Step 1. A measure of information aggregation.** To begin with, we consider the exogenous bounds on the amount of information aggregation. Traders can learn about $v$ *at least* by observing a *single* signal, and *at most* by observing *all* signals. Let $\underline{\tau} < \overline{\tau}$ denote the lower and upper bounds of the precision $\tau$. By Bayes' rule, these bounds are given by

$$\underline{\tau} = \tau_v + \tau_\varepsilon \qquad \text{and} \qquad \overline{\tau} = \tau_v + \tau_\varepsilon \frac{n+1}{N_w}.$$

Recall from (5) that $N_w = 1 + n(1-w)$ and hence is decreasing in $w$. As $w$ increases, the scope of information aggregation increases. Hence, the gap between the two bounds, $\overline{\tau} - \underline{\tau} = n\tau_\varepsilon \frac{w}{N_w}$, increases in $w$.[15] Exactly where the equilibrium value of $\tau$ lies depends on how informative equilibrium prices are. In the Appendix, we show that there is a constant $\varphi \in (0,1)$, to be determined in equilibrium, such that

$$\tau(\varphi) = \tau_v + \tau_\varepsilon \frac{1 - \varphi + w\varphi(n+1)}{1 - \varphi + w\varphi N_w} \in (\underline{\tau}, \overline{\tau}). \tag{11}$$

We use $\varphi$ as a measure of information aggregation because (11) increases in $\varphi \in (0,1)$ with

[15] When information is symmetric (i.e., $w = 0$), everyone observes an identical signal and $\underline{\tau} = \overline{\tau} = \tau_v + \tau_\varepsilon$.

$\tau(0) = \underline{\tau}$ and $\tau(1) = \overline{\tau}$.

**Step 2. A naive conjecture of $\varphi$.** Suppose that a given trader has a naive conjecture $\varphi \in (0,1)$ and conjectures that other traders use (3). The Appendix shows that this trader's best response order has coefficients $(\widetilde{\beta}_s, \widetilde{\beta}_e)$, which satisfy

$$\frac{\tau_\varepsilon}{\rho} \frac{\widetilde{\beta}_e}{\widetilde{\beta}_s} = \frac{1 - \left\{1 + w\left(\frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s} N_u - N_w\right)\right\}\varphi}{1 - \varphi}. \tag{12}$$

Recall that we measure the balance of motives by $k \equiv \frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$. Hence, (12) represents the individual trader's best response in terms of the balance of motives, conditional on the naive conjecture $\varphi$. Rewriting (12) in terms of $k$,

$$\widetilde{k} = BR(k;\varphi) \equiv \frac{1 - \left\{1 + wN_u\left(k - \frac{N_w}{N_u}\right)\right\}\varphi}{1 - \varphi}. \tag{13}$$

For a fixed $\varphi \in (0,1)$ and $w > 0$, the mapping $BR(k;\varphi)$ is linear and decreasing in $k$ with a positive intercept

$$BR(0;\varphi) = 1 + wN_w \frac{\varphi}{1 - \varphi}. \tag{14}$$

Therefore, for a given $\varphi$, the equilibrium balance of motives is a unique solution to $k = BR(k;\varphi)$, which we write as

$$k(\varphi) = \frac{1 + (wN_w - 1)\varphi}{1 + (wN_u - 1)\varphi}. \tag{15}$$

**Figure 1** illustrates $BR(k;\varphi)$ and $k(\varphi)$.

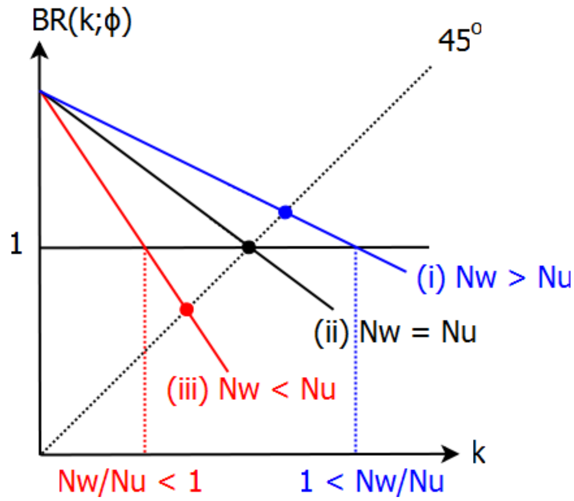


**Figure 1. Best response $BR(k;\varphi)$ for a fixed $\varphi \in (0,1)$ and $w > 0$.**
Note. The higher blue (middle black, lower red) line is the case $N_u < (=, >) N_w$ .

To understand the behavior of $k(\varphi)$, it is helpful to separately analyze the impact of $k$ on $BR(k;\varphi)$ and the impact of $\frac{N_w}{N_u}$ on $BR(k;\varphi)$. The impact of $k$ can be interpreted as an *intensive margin* effect, because it represents the average trader's behavior. The impact of $\frac{N_w}{N_u}$ can be interpreted as an *extensive* margin effect, because it changes with market size $n$.

**Figure 1** shows that $BR(k;\varphi)$ is decreasing in $k$. If the conjectured $k \equiv \frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$ increases (i.e., there is more hedging relative to speculation), the best response $BR(k;\varphi)$ decreases for a given $\varphi$. Therefore, *given that $\varphi$ stays the same*, the balance of motives is *substitutes* among traders. To see why this is the case, we note that two aggregate shocks in prices (10), $\epsilon_0$ and $x_0$, create the secondary roles of $e_i$ and $s_i$ that work *against* their primary roles for hedging and speculation. To begin with, consider $e_i$. The primary role of $e_i$ is to make traders want to reduce the absolute value of $e_i$. This means that hedging intensity $\beta_e$ in (3) must be positive. The secondary role of $e_i$ is its use in the estimation of the common component of endowment shocks $x_0$. When the common shock $x_0$ becomes more important in prices (10) by a conjectured increase in hedging intensity $\beta_e$, prices have a larger predictable component that can be estimated by $e_i$. Because $e_i$ *positively* predicts $x_0$, the greater secondary role of $e_i$ results in a *discount* of hedging intensity. Therefore, other things equal, the best response $\widetilde{\beta}_e$ decreases in the conjectured $\beta_e$. Similar reasoning applies to the behavior of the best response $\widetilde{\beta}_s$ to a conjectured $\beta_s$. Speculative trading is based on the estimation of the payoff value $v$ by $s_i$ (a primary role of $s_i$), hence it requires positive speculation intensity $\beta_s > 0$. An increase in $\beta_s$ in prices (10) makes the estimation of the common component of belief shocks $\epsilon_0$ (a secondary role of $s_i$) more important. Since $s_i$ positively predicts $\epsilon_0$ as well as $v$, the greater secondary role of $s_i$ leads to a discount of speculation intensity. Hence, the best response $\widetilde{\beta}_s$ decreases in the conjectured $\beta_s$. Because an increase in the conjectured $k \equiv \frac{\tau_\varepsilon}{\rho}\frac{\beta_e}{\beta_s}$ raises the importance of the secondary role of $e_i$ *relative to* the secondary role of $s_i$, $\widetilde{\beta}_e$ is discounted by more and $\frac{\widetilde{\beta}_e}{\widetilde{\beta}_s} = BR(k;\varphi)$ decreases in $k$.

Three lines in **Figure 1** represent $BR(k;\varphi)$ for different values of $N_u$ for fixed $\varphi$ and $N_w$, showing that the (negative) slope of $BR(k;\varphi)$ increases in $\frac{N_w}{N_u}$.[16] In other words, the degree of substitution in motives decreases in $\frac{N_w}{N_u}$, which makes $k(\varphi)$ larger. The reason for this is that the relative importance of the non-fundamental role of $e_i$ and $s_i$ depends not only on the conjectured $k$ but also on the variance of $\sqrt{1-w}\epsilon_0 + \sqrt{w}\bar{\epsilon}$ relative to that of $\sqrt{1-u}x_0 + \sqrt{u}\bar{x}$ in prices (10). This is captured by $\frac{N_w}{N_u}$.[17]

Finally, we consider the behavior of $k(\varphi)$. Suppose that $\varphi$ drops for some reason. This results in a lower intercept in **Figure 1**. Because the solution to $BR(k;\varphi) = 1$ is $k = \frac{N_w}{N_u}$ for any $\varphi$, $BR(k;\varphi)$ is lowered for $k < \frac{N_w}{N_u}$ and raised for $k > \frac{N_w}{N_u}$ (i.e., a counter-clock-wise movement around the point $\left(\frac{N_w}{N_u}, 1\right)$). Hence, the change in $\varphi$ has an *asymmetric* impact on $k(\varphi)$ between the two cases: $k(\varphi) \in \left(1, \frac{N_w}{N_u}\right)$ decreases in $\varphi$ ((i) in **Figure 1**, the higher blue dot) and $k(\varphi) \in \left(\frac{N_w}{N_u}, 1\right)$ increases in $\varphi$ ((iii) in **Figure 1**, the lower red dot). The formula for $BR(k;\varphi)$, (13), shows that the direction of the effect of change in $\varphi$ depends on the sign of $k - \frac{N_w}{N_u}$. When $k - \frac{N_w}{N_u} > 0$, $BR(k;\varphi)$ decreases in $\varphi$. To see why this is the case, we first note that for any conjecture of $k$, both $\widetilde{\beta}_e$ and $\widetilde{\beta}_s$ drop in response to an increase in $\varphi$,[18] because greater information aggregation discourages both hedging and speculation.

[16] For a fixed $\varphi$, the intercept of $BR(k;\varphi)$, (14), does not depend on $N_u$, while the slope of $BR(k;\varphi)$, $-\frac{wN_u\varphi}{1-\varphi}$, decreases in $N_u$. As $N_u$ increases relative to $N_w$, $BR(k;\varphi)$ rotates downward around the intercept.

[17] Variance parameters do not show up here because we assumed $Var[x_0] = Var[x_i]$ and $Var[\epsilon_0] = Var[\epsilon_i]$. If the common component and idiosyncratic component of the shock have different variances, the relative variance would matter for the degree of substitution in motives.

[18] We formally prove this claim in the Appendix. See Remark after the proof of **Lemma A1**.

However, the *relative* change in two intensities depends on the overall relative importance of the secondary role of $e_i$ and $s_i$, which depends on $k$ and $\frac{N_w}{N_u}$ through prices (10). Recall that larger $k$ implies smaller $BR(k;\varphi)$ (the intensive margin), while larger $\frac{N_w}{N_u}$ implies the opposite (the extensive margin). When $k-\frac{N_w}{N_u}>0$, the impact of the intensive margin wins. **Lemma 2** summarizes this analysis.

**Lemma 2 (balance of trading motives)**

*Given $w>0$ and fixed $\varphi\in(0,1)$, only one of the following three cases is possible:*

$$\begin{array}{lll} \text{(i) } \textit{If } u > w, & \textit{then } k(\varphi)\in\left(1,\dfrac{N_w}{N_u}\right) & \textit{and } \dfrac{\partial k(\varphi)}{\partial\varphi}>0. \\ \text{(ii) } \textit{If } u = w, & \textit{then } k(\varphi)=1. & \\ \text{(iii) } \textit{If } u < w, & \textit{then } k(\varphi)\in\left(\dfrac{N_w}{N_u},1\right) & \textit{and } \dfrac{\partial k(\varphi)}{\partial\varphi}<0. \end{array}$$

Our analysis so far showed that $k$ is uniquely determined by (15) for a naive conjecture of $\varphi$. We now show however that multiplicity can arise in $\varphi$ itself when traders internalize a systematic link between $\varphi$ and the conjectured $k$.

**Step 3. A rational conjecture of $\varphi$.** A rational conjecture of $\varphi$ must be consistent with the conjectured order (3). From the market-clearing condition (4), information in $p$ from trader $i$'s perspective is summarized by

$$h_i\equiv\frac{n\beta_p p-q_i}{n\beta_s}=v+\sqrt{1-w}\epsilon_0+\frac{\sqrt{w}}{n}\sum_{j\neq i}\epsilon_j-\frac{\beta_e}{\beta_s}\left(\sqrt{1-u}x_0+\frac{\sqrt{u}}{n}\sum_{j\neq i}x_j\right).$$

Note that $[v,s_i,e_i,h_i]^\top$ is jointly normal with mean zero and a variance-covariance matrix that depends on $\frac{\beta_e}{\beta_s}$. In the Appendix we solve the inference problem of $v$ conditional on $(s_i,e_i,h_i)$, and derive the formula for $\tau(\varphi)$ given in (11), but now with $\varphi$ taking a particular value that depends on $\frac{\beta_e}{\beta_s}$. Specifically:

$$\varphi(k) = \left\{1 + \left(\frac{\beta_e}{\beta_s}\right)^2 \frac{\tau_\varepsilon}{\tau_x} u N_u\right\}^{-1} = \left(1 + \alpha_\varepsilon u N_u k^2\right)^{-1} \in (0,1). \tag{16}$$

This shows that the rationally conjectured $\varphi$ *cannot increase* in $k$. This is intuitive because large $k$ implies more hedging and less speculation.[19]

**Lemma 2** indicates the necessary direction of the effect of $k$ on $\varphi$ to create multiplicity. Recall that if $\varphi$ is conjectured independent of the balance of motives $k$, $k$ is a substitute in the sense that $BR(k;\varphi)$ decreases in $k$. To make it a complement, $k$ must affect $\varphi$ in a particular direction. In case (i) in **Lemma 2**, $\varphi$ must increase in $k$ so that a conjectured $k$ can justify a large $BR(k;\varphi)$. But, as shown in (16), this is not possible if traders make a rational conjecture about $\varphi$. By contrast, in case (iii), $\varphi$ must decrease in $k$ to create multiplicity, which is consistent with (16). Therefore, case (iii), i.e., $u < w$, is a *necessary condition for multiplicity*. Intuitively, this condition requires that the secondary role of $e_i$ is large relative to that of $s_i$. if so, traders rationally associate the higher level of information aggregation $\varphi$ with less hedging relative to speculation in equilibrium, i.e., $\frac{\partial k(\varphi)}{\partial \varphi} < 0$. This, combined with $\frac{\partial \varphi(k)}{\partial k} < 0$ from (16), makes multiplicity in $\varphi$ possible.

**Step 4. The role of market size $n$.** To see how large market size is necessary for multiplicity, recall from (5) that we can write

$$\frac{N_w}{N_u} = \frac{\frac{1}{n} + 1 - w}{\frac{1}{n} + 1 - u}$$

[19] Note that the conjecture that $\varphi$ falls in the range of $(0,1)$ is *correct*. An arbitrary value of $\varphi \in (0,1)$ is "naive" only because it is not consistent with (16) derived from the conjectured strategies (3).

and observe that only one of the following three cases is possible:

$$\begin{array}{llll}
\text{(i) } 1 < \dfrac{N_w}{N_u} < \dfrac{1-w}{1-u}, & \dfrac{N_w}{N_u} \text{ increases in } n, & \text{and } \lim\limits_{n\to\infty} \dfrac{N_w}{N_u} = \dfrac{1-w}{1-u}. \\
\text{(ii) } 1 = \dfrac{N_w}{N_u} = \dfrac{1-w}{1-u} & \text{for all } n. & \\
\text{(iii) } 1 > \dfrac{N_w}{N_u} > \dfrac{1-w}{1-u}, & \dfrac{N_w}{N_u} \text{ decreases in } n, & \text{and } \lim\limits_{n\to\infty} \dfrac{N_w}{N_u} = \dfrac{1-w}{1-u}.
\end{array}$$

From **Figure 1**, if $\frac{N_w}{N_u} < 1$ but it is close to 1, the effect of the conjectured $\varphi$ on $k(\varphi)$ is too little to make $k$ a complement. To have a strong impact of $\varphi$ on $k(\varphi)$, it is necessary to have $\frac{N_w}{N_u}$ sufficiently smaller than 1. Since $\frac{N_w}{N_u} > \frac{1-w}{1-u}$ and $\frac{N_w}{N_u}$ monotonically decreases in $n$ in this case, it is necessary to have $\frac{1-w}{1-u}$ sufficiently smaller than 1 and sufficiently large $n$. Intuitively, to make $k$ a complement, the secondary role of $e_i$ must be sufficiently important relative to that of $s_i$. This requires a high correlation between $e_i$ and prices relative to a correlation between $s_i$ and prices. As the number of traders increases, idiosyncratic shocks, $x_i$ and $\epsilon_i$, are averaged out in prices, but common shocks, $x_0$ and $\epsilon_0$, are not. As a result, the gap between $u$ and $w$ becomes more pronounced in larger markets. This can be seen in **Figure 1** as a left-ward movement of the point $\left(\frac{N_w}{N_u}, 1\right)$ for case (iii), which expands the range of $k(\varphi)$ achieved by a change in the conjectured $\varphi \in (0,1)$.

## 3.2 Conditions for multiplicity

The analysis so far showed that it is the joint determination of $k(\varphi)$, (15), and $\varphi(k)$, (16), which makes multiplicity in these variables possible. To derive conditions for multiplicity more precisely, combine (15) and (16) to obtain a cubic equation in $k$:

$$F(k) \equiv \left(\alpha_\varepsilon u N_u k^2 + w\right)(k-1) + \; wn\left\{(1-u)k - (1-w)\right\} = 0. \tag{17}$$

Because $F(k) < 0$ for all $k \leq 0$ and $\lim\limits_{k\to\infty} F(k) = \infty$, this cubic equation has at least one and at most three positive solutions. Therefore, multiple equilibria can be ranked by $k$, and from

$\varphi(k)$, (16), $k$ and $\varphi$ are negatively related across different equilibria. We remind readers that $\alpha_\varepsilon \equiv \frac{\rho^2}{\tau_\varepsilon \tau_x}$ which was given in (5) measures the degree to which the environment is relatively more conductive to hedging than to speculation.

**Proposition 1 (multiplicity)**

*A set of parameters* $(u, w, n, \alpha_\varepsilon)$ *for which multiple equilibria exist is non-empty.*

*In particular, multiplicity arises if the three conditions (a)-(c) all hold:*

(a) *the endowment correlation* $1-u$ *is sufficiently larger than the belief correlation* $1-w$*,*

(b) $n$ *is sufficiently large,*

(c) $\alpha_\varepsilon$ *is neither too small nor too large.*

In the Appendix, we prove that the cubic equation (17) has multiple solutions if and only if the three conditions (a')-(c') all hold:

$$\text{(a')}\ \frac{1-w}{1-u} < \frac{1}{9}, \ \text{ and (b')}\ \frac{8}{1-u-9(1-w)} < n, \ \text{ and (c')}\ \alpha_\varepsilon \in \left[\alpha_\varepsilon^-, \alpha_\varepsilon^+\right], \tag{18}$$

$$\text{where } \alpha_\varepsilon^\pm \text{ are constants which depend on } \frac{w}{u} \text{ and } \frac{N_w}{N_u}.$$

Conditions (a') and (b') simply make the claim already made more precise, i.e., $\frac{1-w}{1-u}$ and $\frac{N_w}{N_u}$ must be sufficiently smaller than 1.[20] Note that condition (a') is necessary for condition (b'). In the Appendix we show that conditions (a') and (b') are necessary and sufficient for the interval $[\alpha_\varepsilon^-, \alpha_\varepsilon^+]$ to exist, and hence they are necessary for condition (c'). To understand condition (c'), note that the rational conjecture of $\varphi$, (16), cannot play a decisive role if $\alpha_\varepsilon$ is either too small or too large.[21] Intuitively, a very small (large) $\alpha_\varepsilon$ makes speculation (hedging) a sufficiently dominant motive, leaving little room for the rationally conjectured $\varphi$ to vary for different conjectures of $k$.[22]

[20] In particular, (b') is equivalent to $\frac{N_w}{N_u} < \frac{1}{9}$.

[21] In the Appendix, we prove that for $w = 1 > u$, $\lim_{n\to\infty} \alpha_\varepsilon^- = \frac{4}{u}$ and $\lim_{n\to\infty} \alpha_\varepsilon^+ = \infty$. In this case, multiplicity is possible for arbitrarily large $\alpha_\varepsilon$. Otherwise, $\lim_{n\to\infty} \alpha_\varepsilon^+ < \infty$.

[22] For example, if traders are risk-neutral, $\alpha_\varepsilon = 0$, hedging is irrelevant and there is no coordination issue.

From **Lemma 1**, an equilibrium exists for sufficiently large $n$. Because condition (a') implies condition (b') for sufficiently large $n$, whether equilibrium multiplicity occurs or not depends on the behavior of the interval $[\alpha_\varepsilon^-, \alpha_\varepsilon^+]$ for large $n$. The following result gives us a theoretical reason why equilibrium multiplicity is more likely to be relevant in large markets.

**Lemma 3 (market size and multiplicity)**

*The set of the value of $\alpha_\varepsilon$ for which multiplicity arises expands with market size $n$.*

We will come back to the implication of a change in market size in the next section, after we characterize the economic properties of multiple equilibria.

# 4 Excess Volatility and Other Properties of Equilibria

To isolate the role of multiplicity for excess volatility, we first investigate excess volatility in a given equilibrium. Second, we study variation in excess volatility and other properties across different equilibria. Finally, we study equilibrium switching induced by a change in market size $n$ or in a market environment captured by $\alpha_\varepsilon$. We discuss empirical implications of the model at the end of the section.

## 4.1 Properties of equilibria

In a linear equilibrium, many properties of the equilibrium can be tied to the coefficients $\left(\beta_s, \beta_e, \beta_p\right)$. Given the orders (3), the equilibrium price and quantities traded $(p, q_i)$ are:

$$p = \frac{\beta_s}{\beta_p}\overline{s} - \frac{\beta_e}{\beta_p}\overline{e} = \frac{\beta_s}{\beta_p}\left(v + \overline{\varepsilon}\right) - \frac{\beta_e}{\beta_p}\overline{e}, \tag{19}$$

$$q_i = \beta_s\left(s_i - \overline{s}\right) - \beta_e\left(e_i - \overline{e}\right), \tag{20}$$

where $\overline{s}$, $\overline{\varepsilon}$, $\overline{e}$, are the market averages of $s_i$, $\varepsilon_i$, $e_i$. We define the following five concepts.

1. *Excess volatility* $Var\left[p|v, \{e_i\}_{i=1}^{n+1}\right]$.

Excess volatility is the variance of prices that is unrelated to the payoff risk $v$ and endowment shocks $e_i$. We take endowment shocks as fundamentals, because otherwise the entire price volatility would become excess volatility in the first-best case of perfect risk-sharing. From (19), the only endogenous variable that affect excess volatility is the ratio $\frac{\beta_s}{\beta_p}$.

2. *Price underreacts to* $v$ *if* $\frac{\beta_s}{\beta_p} < 1$.

If $\frac{\beta_s}{\beta_p} < 1$, price (19) moves less than one-for-one with $v$. This is set against the case $\frac{\beta_s}{\beta_p} = 1$ where price fully reflects the innovation in $v$.

3. *Price impact* $\lambda \equiv \frac{1}{n\beta_p}$.

Price impact measures the extent to which an individual trader moves prices by his own trade. This is a standard notion of illiquidity in the market microstructure literature.[23]

4. *Volume* $\frac{1}{2}E\left[|q_i|\right]$.

Because the quantities traded (20) depend on private information and sum up to zero, volume is measured by the expected absolute value of $q_i$ multiplied by $\frac{1}{2}$. This is again a standard volume measure used in the literature.[24] From the expression of (20), volume depends only on $(\beta_s, \beta_e)$.

5. *Ex ante payoff* $\Pi \equiv -\frac{1}{\rho}\log\left(E\left[\exp\left(-\rho\pi_i\right)\right]\right)$.

Our welfare measure $\Pi$ is the *ex ante* certainty equivalent value of the equilibrium payoff $\pi_i$, defined by $\exp\left(-\rho\Pi\right) = E\left[\exp\left(-\rho\pi_i\right)\right]$. Note that expectation is taken over $(s_i, e_i, p)$. In the Appendix, we show that $\Pi$ is well-defined if $\frac{\rho^2}{\tau_v\tau_x} < 1$.[25] We assume this condition whenever we study $\Pi$.

**Proposition 2 (excess volatility in a given equilibrium)**

*In a given equilibrium, prices underreact to* $v$. *Excess volatility and price impact* $\lambda$ *increase with payoff volatility* $\frac{1}{\tau_v}$, *while volume is independent of* $\frac{1}{\tau_v}$.

[23] See Kyle (1985) and the literature that followed it.
[24] See Vives (2008), p 351.
[25] See Lemma A11 in the Appendix.

**Proposition 2** shows that excess volatility is not driven by the overreaction of prices to the innovation in the payoff value. In the model, excess volatility comes from noise in signals. On the one hand, noise in signals *softens* the reaction of prices to the change in the payoff value, because traders learn the innovation subject to noise, and hence part of it is attributed to noise. On the other hand, when the payoff value is very volatile, traders put more weight on signals, thereby increasing the importance of belief shocks for prices. Thus, the model can jointly explain excess volatility and the underreaction of prices to a change in fundamentals.[26] **Proposition 2** also shows that if variation in excess volatility is driven by variation in payoff volatility, then we should expect excess volatility and price impact moving in the same direction, but no such pattern for volume.

**Proposition 3 (excess volatility across different equilibria)**

*When multiple equilibria exist, an equilibrium with larger $k$ has smaller excess volatility, smaller price impact $\lambda$, larger volume, and larger ex ante payoff $\Pi$.*

For assets subject to multiplicity, we may be observing variation in excess volatility, price impact, and volume across different equilibria. If such is a case, we should expect that excess volatility and price impact move in the same direction, while trade volume should move in the opposite direction. From **Proposition 1**, we can identify the environment where this is likely to be observed: large markets where a correlation in hedging motive is larger than a correlation in speculation motive and neither motive is too dominant.

## 4.2 Multiplicity and equilibrium switching

The Appendix contains an explicit characterization of the interval $[\alpha_\varepsilon^-, \alpha_\varepsilon^+]$, the set of $\alpha_\varepsilon \equiv \frac{\rho^2}{\tau_\varepsilon \tau_x}$ for which multiplicity arises. Since $\alpha_\varepsilon^\pm$ are monotonic in $n$, we define $\underline{\alpha}_\varepsilon \equiv \lim_{n\to\infty} \alpha_\varepsilon^-$ and $\overline{\alpha}_\varepsilon \equiv \lim_{n\to\infty} \alpha_\varepsilon^+$. **Figure 2** illustrates $[\alpha_\varepsilon^-, \alpha_\varepsilon^+]$ for $w < 1$.[27]

[26] The short-run underreaction of prices to fundamental news is well-documented. See Bernard and Thomas (1990), Ikenberry et al (1995), Loughran and Ritter (1995), and Michaely et al (1995).

[27] With $w = 1 > u$, $\underline{\alpha}_\varepsilon = \frac{4}{u}$ and $\overline{\alpha}_\varepsilon = \infty$. See footnote 21.

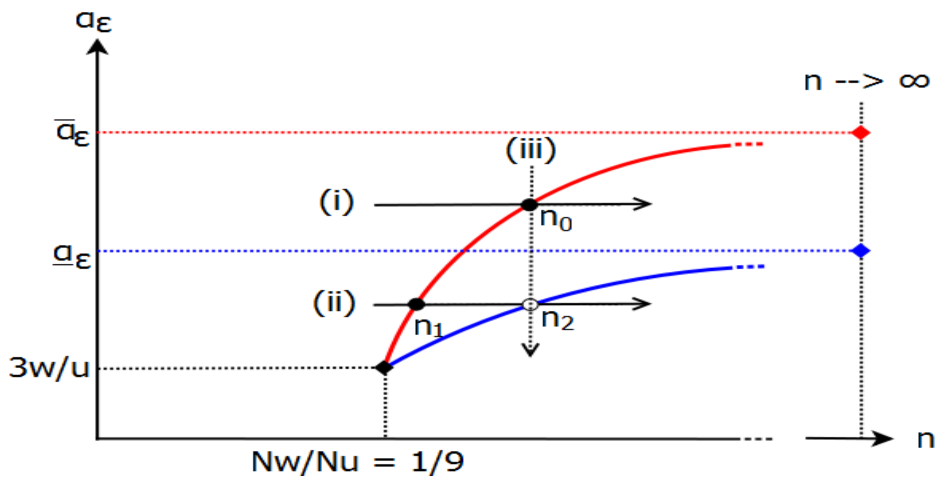


**Figure 2. Multiplicity region.**
Note. The higher red increasing curve is $\alpha_\varepsilon^+$ (lower blue is $\alpha_\varepsilon^-$).

For any given $n$ above the threshold size $\frac{8}{1-u-9(1-w)}$, there is a corresponding interval $[\alpha_\varepsilon^-, \alpha_\varepsilon^+]$ for $\alpha_\varepsilon$ for which multiplicity arises. As $n$ increases, this range expands and shifts upward.[28] We are now ready to state what happens to excess volatility as market size $n$ or an environment captured by $\alpha_\varepsilon$ changes, especially when these parameters reach or leave the multiplicity region (i.e., the three dot points on the two arrows (ii) and (iii) in **Figure 2**).

**Proposition 4 (excess volatility and multiplicity)**

*Whenever multiple equilibria emerge as market size $n$ increases or $\alpha_\varepsilon$ decreases, the new equilibria have larger excess volatility.*

*Whenever equilibria disappear as market size $n$ increases or $\alpha_\varepsilon$ decreases, the equilibrium with the largest excess volatility survives.*

**Remark.** Because each equilibrium corresponds to a solution to the cubic equation (17), on the boundary of the multiplicity region (points $n_0$, $n_1$, $n_2$ in **Figure 2**), (17) has a double root and another root. In the Appendix, we show that $\alpha_\varepsilon = \alpha_\varepsilon^+$ (points $n_0$, $n_1$)

[28]If $\alpha_\varepsilon \in [\underline{\alpha}_\varepsilon, \overline{\alpha}_\varepsilon)$, multiplicity arises for all $n$ above some threshold (see (i) and $n_0$ in **Figure 2**). Combined with **Lemma 1**, we provided a strategic foundation to multiplicity in large markets for the parameter region $\frac{1-w}{1-u} < \frac{1}{9}$ and $\alpha_\varepsilon \in [\underline{\alpha}_\varepsilon, \overline{\alpha}_\varepsilon)$. If $\alpha_\varepsilon = \overline{\alpha}_\varepsilon$, multiplicity arises only in the limit $n \to \infty$ but not for any finite $n$. If $\alpha_\varepsilon \in \left(3\frac{w}{u}, \underline{\alpha}_\varepsilon\right)$, multiplicity can arise only for finite $n$. See (ii) and $n_1$, $n_2$ in **Figure 2**.

corresponds to the case where a double root is smaller than the other solution, while $\alpha_\varepsilon = \alpha_\varepsilon^-$ (point $n_2$) corresponds to the case where a double root is larger than the other solution.

**Proposition 4** indicates that whenever an equilibrium switch occurs due to an increase in market size $n$, either because new equilibria emerge or the current equilibrium disappears, *excess volatility must jump up*. Similarly, an equilibrium switch due to a decrease in $\alpha_\varepsilon \equiv \frac{\rho^2}{\tau_\varepsilon \tau_x}$ results in a jump-up in excess volatility. Recall that a decrease in $\alpha_\varepsilon$ can be interpreted as the underlying environment becoming more conductive to speculation. Thus, in the parameter region where multiplicity is possible, equilibrium excess volatility can be quite sensitive to market size $n$ as well as to the market environment as captured by $\alpha_\varepsilon$. Note that the reverse change can also occur: an equilibrium switch due to a decrease in $n$ or an increase in $\alpha_\varepsilon$ results in a jump-*down* in excess volatility. Hence, the model predicts a particular direction of a jump in excess volatility depending on the nature of the change in the environment.

**Empirical implications.** The model explains variation in excess volatility in two ways. First, an asset with higher payoff volatility $\frac{1}{\tau_v}$ has higher excess volatility because noisy signals will be given more weight in speculation. Second, in large markets, excess volatility can be sensitive to a small change in market size $n$ or a change in the environment (i.e., a change in $u$, $w$, $\alpha_\varepsilon$).[29] These mechanisms yield the following testable implications.

- Across different assets ordered by their payoff volatility $\frac{1}{\tau_v}$, or for a single asset with time-varying payoff volatility, the model predicts a comovement of payoff volatility, excess volatility, and price impact, but not trade volume.

- For assets subject to multiplicity, the model predicts that a sudden and large change in excess volatility is possible in response to a small change in the trading environment.

- When the above occurs, excess volatility and price impact should move together while volume should move in the opposite direction.

[29] Interestingly, the second mechanism is completely orthogonal to the first mechanism, because regardless of the level of payoff volatility of the asset, the same condition for equilibrium multiplicity applies.

- A sudden increase in excess volatility is more likely when assets attract more traders, or when the environment becomes more conductive to speculation. A sudden decrease in excess volatility is more likely for the opposite cases.

One way to test a part of this theory based on multiplicity is to look for two assets that are similar in their payoff volatility, but where only one of them satisfies the conditions for multiplicity. The model predicts that only one asset should exhibit a sudden and large change in excess volatility, and the systematic relationship between excess volatility, price impact, and volume due to equilibrium switching. One candidate of such pair of assets is the closed-end fund and its underlying portfolio of assets. By construction, their payoff volatilities are identical, but the closed-end fund is likely to have more traders.[30] In fact, Pontif (1997) studied excess volatility in this context, and found that the average closed-end fund's monthly return was 64% more volatile than its assets. He also found that prices underreacted to fundamentals, and excess volatility was largely (85%) idiosyncratic rather than market-wide. These findings are consistent with the model.[31] If excess volatility of the closed-end fund is indeed generated by equilibrium multiplicity, the model additionally predicts a particular comovement of excess volatility, price impact, and volume. We leave the empirical investigation of the theory for future research.

**Related literature**. Multiplicity of the same nature has been studied by Ganguli and Yang (2009) and Manzano and Vives (2011). We depart from their continuum-traders framework by studying a game among a finite number of traders.[32] This allows us to study the connection between market size and multiplicity, and other equilibrium properties. Also, while they focused on strategic complementarity/substitutability in traders' actions and in

[30] Of course, there are many other dimensions in which they can be different (e.g. type of traders attracted to each asset). Here we are discussing excess volatility after controlling for other relevant differences.

[31] The source of excess volatility in our model is signal noise. This is likely to be asset specific. For example, idiosyncratic aspects of the fund, such as its name, may create a common shock in buyers' beliefs.

[32] Diamond and Verecchia (1981) first introduced multiple trading motives in a model with a finite number of traders. They used a price-taking equilibrium as a solution concept, but we show that multiplicity does not occur in their environment even if strategic behaviors were taken into consideration.

information acquisition, we focus on excess volatility, price impact, volume, and welfare. We believe that these are equally important properties to be compared across equilibria.

Excess volatility has been extensively documented in the finance literature, but its origin is largely an unresolved issue. Timmermann (1993) and others[33] proposed the estimation uncertainty of model parameters as an explanation. We view our multiplicity-based explanation as complementary to the estimation uncertainty-based explanation, because adding the need to estimate model parameters in the presence of equilibrium multiplicity would increase estimation uncertainty. This also strengthens our argument that excess volatility can increase with market size.

The model in this paper can be interpreted as a model of rational bubbles, where volatility of a bubble component increases with fundamental volatility. Because the model provides an endogenous link between fundamental and nonfundamental components in prices, it shares the similarity with a model of intrinsic bubbles by Froot and Obstfeld (1991). The key difference is that in their model prices overreact to the fundamental innovation while in the current model prices underreact to it.

There is a large literature studying multiplicity as a source of nonfundamental volatility.[34] We contribute to this literature by offering a theoretical reason why market size may be relevant for multiplicity. According to our model, large markets with highly sophisticated traders may be prone to multiplicity.

# 5 Conclusion

We studied excess volatility and equilibrium multiplicity in a model of trading with a finite number of traders. The multiplicity is possible only in large markets where a rational inference from prices makes traders' motives complements. We showed that multiple equilibria

[33] See also Barsky and De Long (1993).

[34] For example, Angeletos and Werning (2006) studied a role of endogenous information and nonfundamental volatility by introducing a financial market into a standard coordination game. In their model, multiplicity emerges in prices but neither in individual strategies nor in volume.

are ranked in terms of excess volatility, price impact, volume, and welfare. The model suggests that when traders have multiple trading motives, a coordination issue can arise and the failure to coordinate on fundamental trading may lead to a high level non-fundamental volatility. In our model, multiplicity arises precisely because *many* traders simultaneously conduct *sophisticated* inferences from prices. As such, we view the identified mechanism as relevant for extreme volatility observed in markets with a large number of sophisticated traders. Our model also suggests a potential cost of organizing large financial markets.

# References

[1] Angeletos, G. M., and I. Werning (2006): "Crises and Prices: Information Aggregation, Multiplicity, and Volatility." *American Economic Review*, 96, 5, 1720-1736.

[2] Barsky, B., and J. B. De Long (1993): "Why Does the Stock Market Fluctuate?" *Quarterly Journal of Economics*, 108, 2, 291-311.

[3] Bernard, V. L., and J. K. Thomas (1990): "Evidence That Stock Prices Do Not Fully Reflect the Implications of Current Earnings for Future Earnings." *Journal of Accounting and Economics*, 13, 305-340.

[4] Diamond, D. W., and R. E. Verrecchia (1981): "Information Aggregation in a Noisy Rational Expectations Economy." *Journal of Financial Economics*, 9, 221-235.

[5] Froot, K. A. and M. Obstfeld (1991): "Intrinsic Bubbles: The Case of Stock Prices." *American Economic Review*, 81, 5, 1189-1214.

[6] Gabaix, X., P. Gopikrishnan, V. Plerou, and H. E. Stanley (2006): "Institutional Investors and Stock Market Volatility" *Quarterly Journal of Economics*, 121, 2, 461-504.

[7] Ganguli, J. V., and L. Yang (2009): "Complementarities, Multiplicity, and Supply Information." *Journal of the European Economic Association*, 7, 1, 90-115.

[8] Ikenberry, D., J. Lakonishok, and T. Vermaelen (1995): "Market Underreaction to Open Market Share Repurchases." *Journal of Financial Economics*, 39, 181-208.

[9] Kyle, A. S. (1985): "Continuous Auctions and Insider Trading." *Econometrica*, 53, 6, 1315-1335.

[10] Kyle, A. S. (1989): "Informed Speculation with Imperfect Competition." *Review of Economic Studies*, 56, 317-356.

[11] Loughran, T. , and J. R. Ritter (1995): "The New Issues Puzzle." *Journal of Finance*, 50, 1, 23-51.

[12] Michaely, R., R. H. Thaler, and K. L. Womack (1995): "Price Reactions to Dividend Initiations and Omissions: Overreaction or Drift?" *Journal of Finance*, 50, 2, 573-608.

[13] Manzano, C., and X. Vives (2011): "Public and Private Learning from Prices, Strategic Substitutability and Complementarity, and Equilibrium Multiplicity." *Journal of Mathematical Economics*, 47, 346–369.

[14] Pontiff, W. (1997): "Excess Volatility and Closed-End Funds," *American Economic Review*, 87, 1, 155-169.

[15] Rostek, M. and M. Weretka (2012): "Price Inference in Small Markets," *Econometrica*, 80, 2, 687-711.

[16] Shiller, R. J. (2003): "From Efficient Markets Theory to Behavioral Finance," *Journal of Economic Perspective*, 17, 1, 83-104.

[17] Shiller, R. J. (2014): "Speculative Asset Prices," *American Economic Review*, 104, 6, 1486-1517.

[18] Timmermann, A. G. (1993): "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *Quarterly Journal of Economics*, 108, 4, 1135-1145.

[19] Vives, X. (2008): "Information and Learning in Markets." *Princeton University Press.*

[20] Vives, X. (2011): "Strategic Supply Function Competition with Private Information," *Econometrica*, 79, 6, 1919-1966.

[21] West, K. D. (1988): "Dividend Innovations and Stock Price Volatility," *Econometrica*, 56, 1, 37-61.