Research article

Intelligence and global bias in the stock market

Kazuo Sano*

SOBIN Institute LLC, 3-38-7 Keyakizaka, Kawanishi, Hyogo 666-0145, Japan

* **Correspondence:** sano@sobin.org.

Abstract: Trade is one of the essential features of human intelligence. The securities market is the ultimate expression of it. The fundamental indicators of stocks include information about the effects of noise and bias on stock prices; however, distinguishing between them is generally hard. In this article, I present the fundamentals hypothesis based on rational expectations and detect the global bias components from the actual fundamental indicators using a log-normal distribution model based on the fundamentals hypothesis. The analysis results show that biases generally exhibit the same characteristics, strongly supporting our theory. Notably, the positive price-to-cash flows from the investing activities ratio are proxies for the fundamentals. The answer is simple: "Cash is a fact, and profit is an opinion." Namely, opinions of management and accounting added noise to fundamentals. As a result, we obtain the Kesten process and the Pareto distribution. This result means the market knows this noise and shows a stable global bias in the stock market.

Keywords: stock price; noise; bias; log-normal distribution; Gibrat's law; Kesten process

JEL Codes: G14, G32, M40, M41

1. Introduction

Can cats or dogs barter? Trade is one of the essential features of human intelligence. Brosnan et al., (2008) reported that chimpanzees, who are an intelligent species like humans, can trade but are reluctant to trade. This subsequently led to major divergence in the fates of the two species. The market is nothing but an expression of human intelligence. As Malone and Bernstein (2015) mentioned, intelligence does not arise only in individual brains; it also arises in groups of individuals. The securities market, including crypto-assets, is the ultimate expression of human intelligence. Fama (1965) stated that the efficiency of the securities market in the strong or semi-strong sense causes instantaneous changes in traders' subjective equilibrium with the incorporation of securities information, and the speed at which they respond to this information determines the winners and losers. In the weak sense, market

equilibrium follows changes in traders' subjective equilibria due to noise. The efficiency of the securities market can be summarized as the correct and rapid incorporation of information into prices, the presence of arbitrage opportunities, countless traders always searching for such opportunities, the gradual loss of arbitrage opportunities, and the market becoming more efficient. However, it should be noted that without noise, transactions would not occur, and the price distortion would persist in the presence of strong bias.

In an efficient market where noise and bias have no effect, and information is perfectly symmetrical, security prices should accurately reflect only information. However, if traders had perfectly and simultaneously symmetrical knowledge and information, including the asset valuation model, the transaction would not function because the traders' valuation of the asset would be the same. They must be doing some rational calculations via intelligence in natural and digital computing. In this sense, a rational representative agent in macroeconomics corresponds to the perfect symmetry of information (Lucas (1976), Kirman (1992), Hartley (1996)).

While Black (1986) treated this "symmetry breaking" as noise, the effect of noise on a security's price is expected to be symmetrical based on its nature. However if there exist so many irrational noise traders synchronizing erroneous stochastic beliefs that both affect prices and earn higher expected returns, the unpredictability of noise traders' beliefs creates a risk in the price of the asset that deters rational arbitrageurs from aggressively betting against them. As a result, prices can diverge significantly from fundamental values even in the absence of fundamental risk. Moreover, bearing a disproportionate amount of risk that they themselves create enables noise traders to earn a higher expected return than rational investors (De Long et al., (1990)).

On the other hand, Tversky and Kahneman (1974) found that the effect of bias is asymmetric. They described three heuristics that are employed in making judgments under uncertainty: 1) representativeness, 2) availability of instances or scenarios, and 3) adjustment from an anchor. These heuristics are highly economical and usually effective, but they lead to systematic and predictable errors. However, the effect of these heuristics has not been detected as a global bias in the securities market. Consequently, identifying the specific effects of noise and bias on security prices is challenging.

If attention is paid to any statistical property in any complex system, the log-normal distribution is the most natural and appropriate among the standard or "normal" statistics to overview the whole system (Kobayashi et al., (2011)). Log-normality emerges as a familiar and typical example of statistical aspects in various complex systems. Since every member of any complex system has its own history, each member is in the process of growth (or retrogression). The log-normal distribution is realized because of Gibrat's law, or the Matthew effect. It is applied to cities' size and growth rate, where the proportionate growth process may cause a distribution of city sizes that is log-normal. When considering the entire size distribution, not just the largest cities, the city size distribution is log-normal (Samuels (1965)). However, it has been argued that it is problematic to define cities through their fairly arbitrary legal boundaries. According to Gabaix (1999), Zipf's law is a very tight constraint on the class of admissible models of local growth. It says that for most countries, the size distribution of cities strikingly fits a power law: the number of cities with populations greater than S is proportional to 1/S. Suppose that, at least in the upper tail, all cities follow some proportional growth process (this appears to be verified empirically). This automatically leads their distribution to converge to Zipf's law.

According to Gibrat's law of proportionate effect, the relative change in a firm's size is not affected by its absolute size, which implies that both small and large firms experience similar average rates of

growth. However, Samuels (1965) provides evidence that contradicts this law by showing that larger firms grow significantly faster. Aoki and Nirei (2017) developed a neoclassical growth model that produces Pareto's law of income distribution and Zipf's law of firm size distribution based on firm-level productivity shocks. Executives and entrepreneurs invest in both risk-free assets and their own firms' risky stocks, which affect their income and wealth. They used the model to investigate how changes in tax rates can explain the evolution of top incomes in the US. The model accounts for the recent decline in the Pareto exponent of the income distribution and the trend of the top 1 percent income share. In the same research direction, Nirei and Aoki (2016) developed a neoclassical growth model that incorporates heterogeneous households and explains the Pareto distributions of income and wealth in the upper tail. They introduced households' business productivity risks and borrowing constraints in a standard Bewley (1977) model to generate Pareto distributions. Low-productivity households rely on wages and safe asset returns, whereas high-productivity households do not diversify their business risks. The model can accurately explain the observed income distribution in the United States with reasonable parameter values. The authors conducted comparative statics to investigate how changes in parameters affect Pareto distributions and found that the increase in top income dispersion in recent decades may be attributed to changes in the top tax rates in the 1980s. Their analytical finding offers a consistent explanation for numerical comparative statics.

Few studies are available on global bias in the stock market. The fundamental indicators of stocks include information about the effects of noise and bias on stock prices; however, distinguishing between them is generally hard. In this article, I present the fundamentals hypothesis based on rational expectations (Muth (1961)) and, using a log-normal distribution model, detect global bias components from the price-earnings (P/E), price-to-book (P/B) and price-to-cash flow (P/CF) ratios. The traditional theory of the firm assumes that the firm acts in the stockholders' interests and that stockholders are interested in profits such that the object of the firm is to maximize profit. However, there is a range in the profit concept (Bodenhorn (1964)). The analysis results support our fundamentals hypothesis as the detected biases show similar characteristics. Additionally, the results show that the cash flow indicators contain relatively few bias components and are closer to the fundamentals. I further demonstrate and examine why the positive P/IC ratio among the indicators analyzed is a proxy for the fundamentals that do not include bias components. Globally, this implies that the efficient market hypothesis holds. The shape of the curve, and in particular the strength of the bias, is stable throughout time and independent of whether the economy is good or bad. The answer is simple: Cash is a fact, and profit is an opinion. Namely, opinions of management and accountants are added as noise to fundamentals.

2. Materials and method

2.1. Hypothesis

When X_t denotes the true fundamentals of listed companies at time t, and R_t denotes their growth rates, the following Gibrat process represents the growth of those companies

$$X_t = R_t X_{t-1}. \tag{1}$$

It is important to note that I assume that the growth rates R_t are mutually independent random variables that follow the same distribution with finite variance. The initial value of the fundamentals being set as

X_0 yields

$$X_T = X_0 \prod_{t=1}^{T} R_t \tag{2}$$

at time T. Taking the log of both sides of the equation results in

$$\log X_T = \log X_0 + \log R_1 + \cdots + \log R_T. \tag{3}$$

Therefore,

$$\log X_T \sim LN(\mu, \sigma^2) \tag{4}$$

would hold for a sufficiently large T based on the central limit theorem. Essentially, the true fundamental X_T of listed companies follows the log-normal distribution.

Furthermore, by assuming rational expectations through the future point in time T as of the present point in time 0 on the premise of a going concern, the following equation becomes true:

$$X_0 = \mathbb{E}[X_T] \prod_{t=1}^{T} \mathbb{E}[R_t^{-1}]. \tag{5}$$

Therefore, the rational expectations X_0 for fundamentals follow the log-normal distribution. In other words, the fundamentals X_t at time t follow the log-normal distribution.

2.2. Data

Among the actual fundamental indicators, we analyze the price-earnings (P/E), forward price-earnings (P/FE), price-to-book (P/B), price-to-cash flows from operating activities (P/OC), price-to-cash flows from investing activities (P/IC) price-to-cash flows from financing activities (P/FC), and the price-to-cash equivalents at the year-end (P/CE) ratios. We divide the cash flow ratios into positive and negative data. We use daily data with absolute values less than 1,000 for all companies listed in Japan for the period spanning 1,817 business days from January 2007 to May 2014 (data source: Nikkei NEEDS https://nkbb.nikkei.co.jp/en/service/nikkei-needs/).

2.3. Test

If our fundamentals hypothesis is true, the actual fundamental indicators should follow the log-normal distribution. To test our hypothesis, we calculate the p-values using the Kolmogorov-Smirnov test, Pearson's χ^2 test and Anderson-Darling test, to test the goodness of fit to the log-normal distribution.

Our tests can be strongly affected by the sample size that differs by indicator, so we extract samples at the 300 quartiles from each data set to calculate the p-value for testing the goodness of fit of the overall average of each indicator to the log-normal distribution with variance. For all cases, the null hypothesis is that "the indicator follows the log-normal distribution," or the otherwise worded "the indicator is a proxy of true fundamentals." The indicator reflects the fundamentals more when the p-value is closer to 1, and when p-value is closer to 0, the bias is stronger. We show the lowest p-value for each test in Table 1 for the data spanning 1,817 business days.
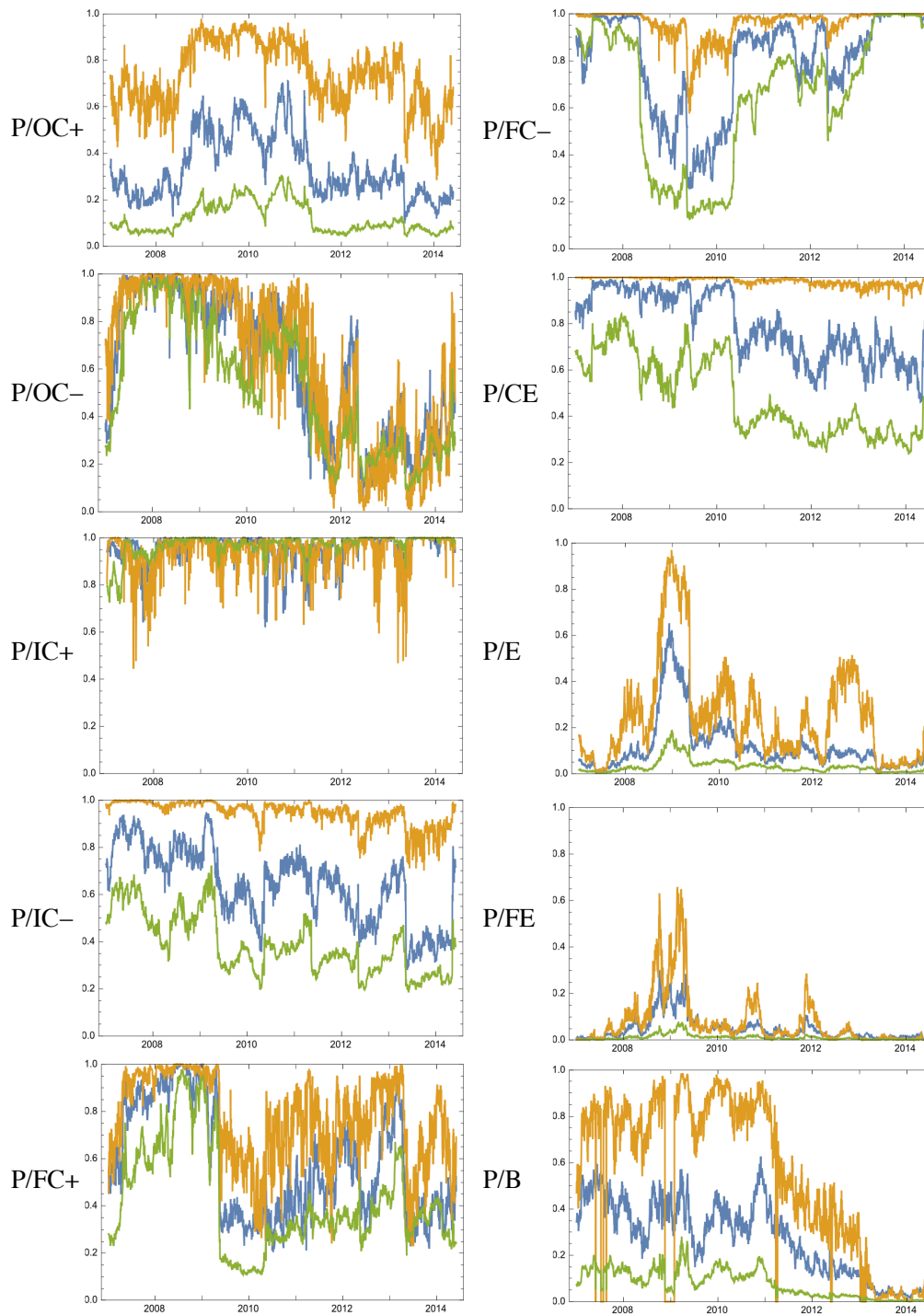
Figure 1. Time series of the goodness of fit test to the log-normal distribution. Blue: Kolmogorov-Smirnov, Yellow: Pearson's χ^2, Blue: Anderson-Darling. The higher the value, the less bias there is, and vice versa.
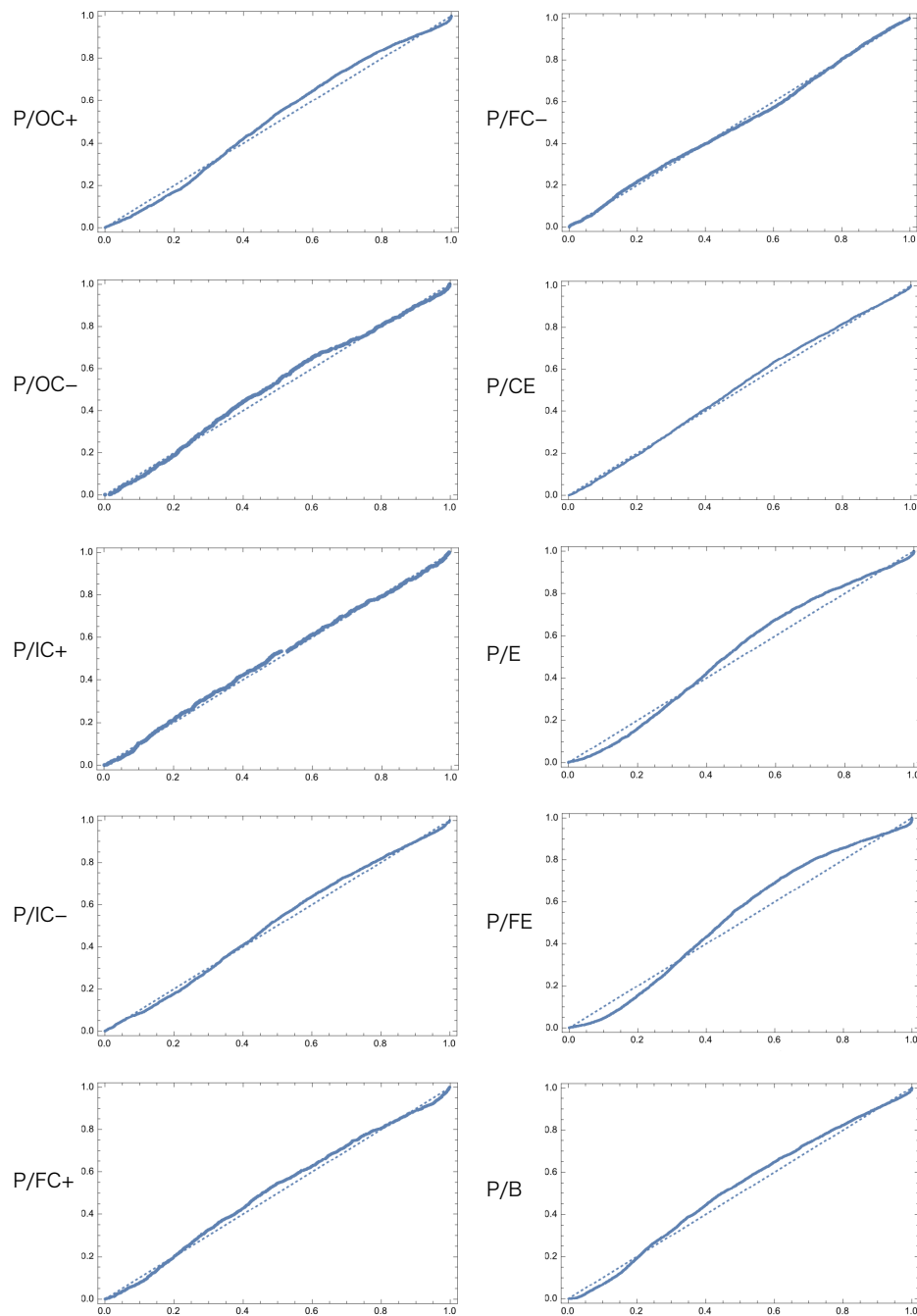
Figure 2. The more the data deviates from a straight line showing a log-normal distribution, the stronger the bias. These graphs are drawn with data from the first day of the time series. The shape of the bias varies proportionally to Figure 1, but is stable throughout the time. `https://figshare.com/projects/Intelligence_and_Global_Bias_in_the_Stock_Market/131711`

The test results indicate that bias is strong on the P/E, P/FE, and P/B ratios. The forward price-earnings ratio (P/FE) is the most biased. Among all price-to-cash flow ratios, the bias is significant only for the P/OC ratio. It can be easier to manipulate than others. I illustrate the time series of p-values since the p-value is the indicator of the strength of bias. From Table 1, Figure 1, and Figure 2, the positive P/IC ratio is the best proxy for the fundamentals that do not include bias components. Probably, it should be hard to manipulate according to the opinions of management and accounting.

Table 1. The lowest p-values for the data spanning 1,817 business days. Samples at the 300 quartiles. Significant (bias) levels: 1%(***), 3%(**), 5%(*)

	Kolmogorov-Smirnov	**Pearson's χ^2**	**Anderson-Darling**
P/OC+	0.0986	0.2843	0.0398*
P/OC−	0.0634	0.0045***	0.0850
P/IC+	0.6205	0.4441	0.7201
P/IC−	0.2787	0.7012	0.1849
P/FC+	0.2042	0.2287	0.1093
P/FC−	0.2522	0.5763	0.1269
P/CE	0.4611	0.8740	0.2376
P/E	0.0170**	0.0000***	0.0018***
P/FE	0.0107**	0.0040***	0.0019***
P/B	0.0032***	0.0023***	0.0005***

2.4. Shape of Bias

The detected biases exhibit similar characteristics. Each indicator is shown in Figure 2 and compared to the log-normal distribution. While the figures only show the data for the first day of all periods, January 4, 2007, the biases remain unchanged. It is worth noting that the shape and strength of the curve are independent of whether the economy is good or bad. The artificial stock market appears to have natural intelligence.

3. Discussion

The results indicate that the fundamental indicators, including the P/E, P/FE, and P/B ratios, are strongly affected by bias. Bias also has a significant effect on the positive and negative P/OC ratios. Additionally, there is a weak bias on the negative P/IC, positive and negative P/FC, and P/CE ratios.

When we compare the test results, the positive P/IC ratio is the stable proxy of fundamentals among all these indicators. Positive cash flows from investing activities represent the realized gain or loss from past investments such as marketable securities, tangible fixed assets, the sales of investment securities, and income from the collection on loans declared at the end of the accounting period. In other words, positive cash flow is the indicator that most directly reflects past business decisions. We can interpret the year-end cash equivalent ratio and the other cash flow ratios reflecting the fundamentals because they have less bias of the components.

On the other hand, although the P/E ratio and P/B ratio are definite values, investors might not view them as indicators that reflect the fundamentals because there is a high degree of freedom in accounting.

The P/OC ratio might also have a lower credibility than other cash flow indicators.

The answer is simple: "Cash is a fact, and profit is an opinion." Namely, opinions of management and accountants are added as noise to true fundamentals. As a result, the Kesten (1973) process

$$X_t = R_t X_{t-1} + \epsilon_t, \quad E[\epsilon_t] > 0, \tag{6}$$

is realized so that the Pareto distribution is obtained. This result means that a positive bias accompanied their opinions.

In fact, these biases fit the Pareto distribution quite well. The following functions represent generalized Pareto distribution (GPD).

$$F(x) = 1 - \left(1 + \left(\frac{x-\mu}{\kappa}\right)^{1/\gamma}\right)^{-\alpha} \tag{7}$$

$$f(x) = \frac{\alpha}{\gamma}\kappa^{-1/\gamma}(x-\mu)^{-1+\frac{1}{\gamma}}\left(1 + \left(\frac{x-\mu}{\kappa}\right)^{1/\gamma}\right)^{-\alpha-1} \tag{8}$$
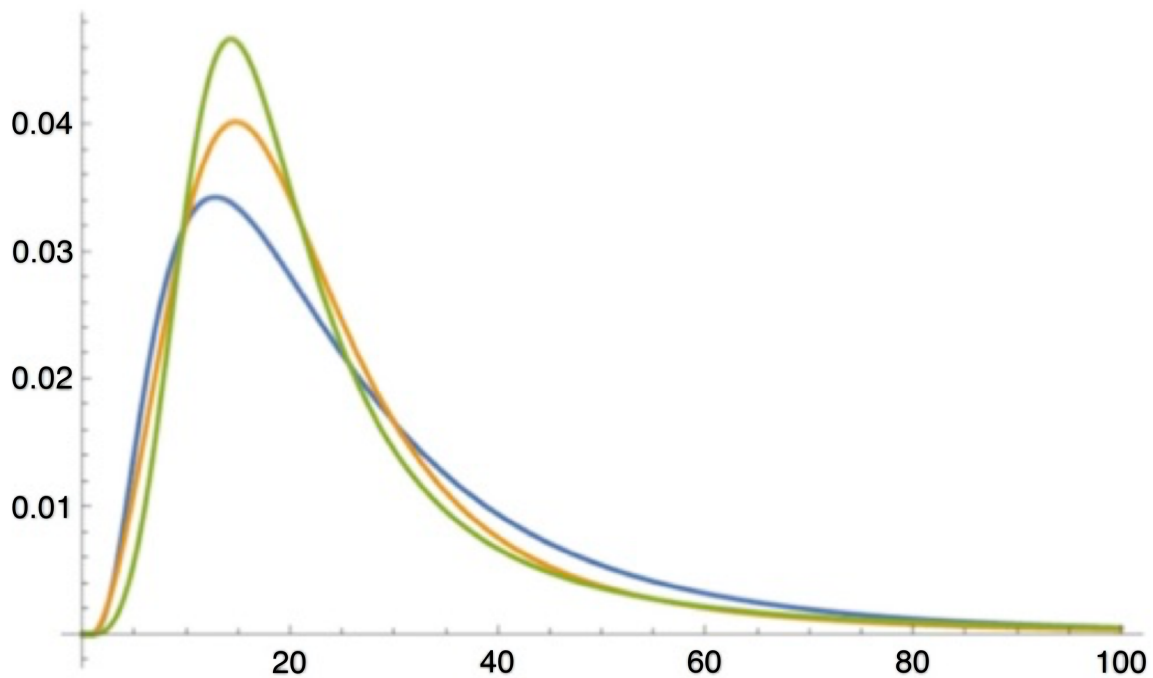

Figure 3. Comparison of three distributions: Log-Normal (blue), GDP-1 (yellow), GDP-2 (green)
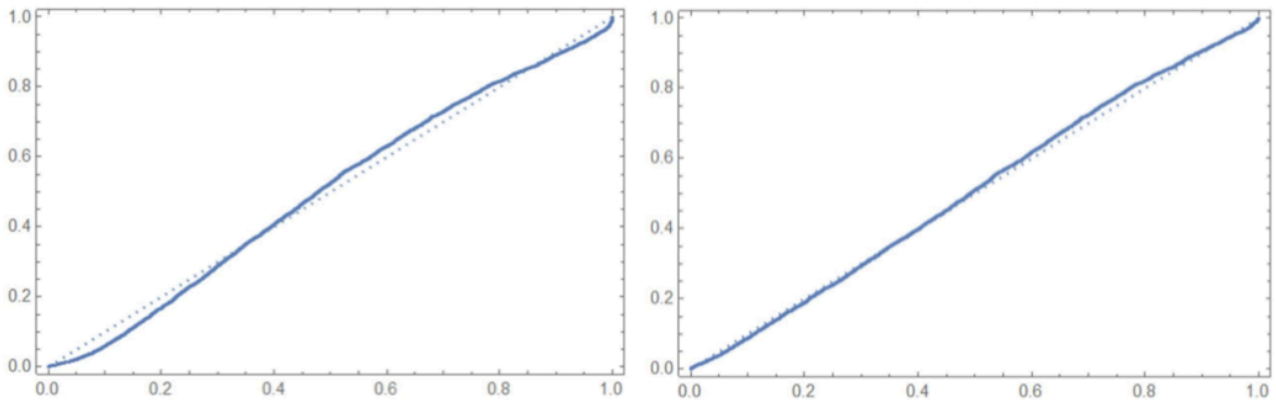
Figure 4. Left: GDP-1 to fit P/FE ($\kappa = 18.82, \alpha = 1, \gamma = 0.385, \mu = 0.993$), Right: GDP-2 to fit P/FE ($\kappa = 13.70, \alpha = 0.515, \gamma = 0.238, \mu = 0.993$). Note that P/FE has the strongest bias.
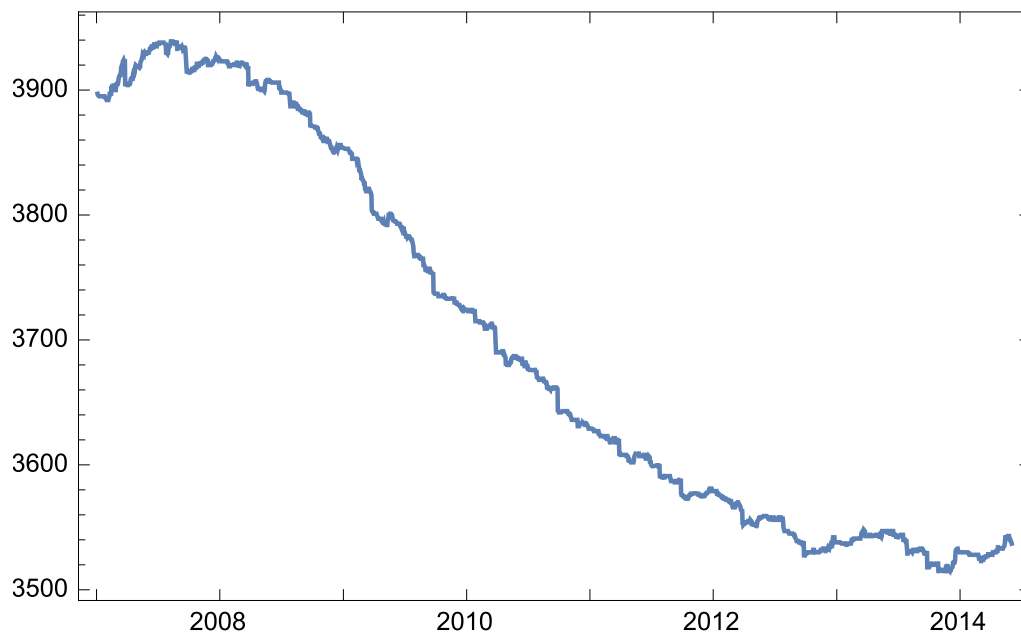

Figure 5. Number of all listed companies in the Japanese stock market (data source: Nikkei NEEDS https://nkbb.nikkei.co.jp/en/service/nikkei-needs/)

cash flow statement

Cash flow from operating activities
Net income before taxes +
Depreciation and amortization +
Increase in accounts receivable -
Increase in inventory -
Increase in accounts payable +
Income taxes paid -
(Total → ①)

Cash flow from investing activities
Purchase of tangible (intangible)
assets -
Sale of tangible (intangible) assets +
Purchase of securities -
Sale of securities and maturity
  redemption +
(Total → ②)

Cash flow from financing activities
Increase in current or long-term debt
+
Dividend payments -
Repayment of current or long-term
debt -
(Total → ③)

Increase in cash and cash equivalents
① + ② + ③= ④

**Cash and cash equivalents at
beginning of year ⑤**

**Cash and cash equivalents at the end
of the period ④ + ⑤**

- 2007/1～2014/5（Nikkei NEEDS）
- 3515～3939 listed companies
- price-operating cash flow ratio（P/OC±）
- price-investing cash flow ratio（P/IC±）
- price-financial cash flow ratio（P/FC±）
- price-cash equivalents ratio（P/CE）
- price-earnings ratio（P/E）
- expected price-earnings ratio（P/FE）
- price-book value ratio（P/B）

Data source: https://nkbb.nikkei.co.jp/en/service/nikkei-needs/

4. Conclusions

What are the implications of the results? Does the existence of bias negate the efficient market hypothesis? From the viewpoint of intelligence in natural and digital computing, how should we deal with this problem? The artificial stock market, where countless traders use digital computers to trade, appears to have natural intelligence. Intelligence does not arise only in individual brains but also arises in groups of individuals. The interest in whether markets are efficient is based on the idea that inefficiencies are the reason active returns are possible. If our fundamentals hypothesis is true, the fundamental indicators, including the P/E, P/FE, and P/B ratios, are strongly affected by bias. This result means that investors might not view them as indicators that reflect the fundamentals since they have a high degree of freedom in accounting. Can we conclude that the market is inefficient because of this? This issue has important economic implications. Globally, our result implies that the efficient market hypothesis holds. The shape of the curve, and in particular the strength of the bias, is stable throughout time and independent of whether the economy is good or bad. It is the path of the ants (Kirman (1993); Sano (2022)) to avoid obstacles, and traders move back and forth along it in response to noise (Black (1986); De Long et al., (1990)). This noise, among other things, plays a role in activating the market system, which knows the fundamentals of firms and represents a stable global bias in the stock market. Nevertheless, when the biases revealed by this study disappear, the stock market will become more efficient. It is the responsibility of firms and financial authorities to disclose fair information.

Use of AI tools declaration

The author declares he has not used Artificial Intelligence (AI) tools in the creation of this article.

Conflicts of interest

The author declares no conflicts of interest in this paper.

References

Aoki S and Nirei M (2017) Zipf's law, pareto's law, and the evolution of top incomes in the united states. *Am Econ J-Macroecon* 9: 36–71. https://doi.org/10.1257/mac.20150051

Bewley T (1977) The permanent income hypothesis: A theoretical formulation. *J Econ Theory* 16: 252–292. https://doi.org/10.1016/0022-0531(77)90009-6

Black F (1986) Noise. *J Financ* 41: 528–543. https://doi.org/10.1111/j.1540-6261.1986.tb04513.x

Bodenhorn D (1964) A cash-flow concept of profit. *J Financ* 19: 16–31. https://doi.org/10.2307/2977477

Brosnan SF, Grady MF, Lambeth SP, et al. (2008) Chimpanzee autarky. *PLOS ONE* 3: 1–5. https://doi.org/10.1371/journal.pone.0001518

De Long JB , Shleifer A, Summers LH, et al. (1990) Noise trader risk in financial markets. *J Polit Econ* 98: 703–738. https://doi.org/10.1086/261703

Fama EF (1965) The behavior of stock-market prices. *J Bus* 38: 34–105. https://doi.org/10.1086/294743

Gabaix X (1999) Zipf's Law for Cities: An Explanation. *Q J Econ* 114: 739–767. https://doi.org/10.1162/003355399556133

Hartley JE (1996) Retrospectives: The origins of the representative agent. *J Econ Perspect* 10: 169–177. https://doi.org/10.1257/jep.10.2.169

Kesten H (1973) Random difference equations and Renewal theory for products of random matrices. *Acta Math* 131: 207–248. https://doi.org/10.1007/BF02392040

Kirman AP (1992) Whom or what does the representative individual represent? *J Econ Perspect* 6: 117–136. https://doi.org/10.1257/jep.6.2.117

Kirman A (1993) Ants, rationality, and recruitment. *Q J Econ* 108: 137–156. https://doi.org/10.2307/2118498

Kobayashi N, Kuninaka H, Wakita J, et al. (2011) Statistical features of complex systems: toward establishing sociological physics. *J Phys Soc Japan* 80: 072001. https://doi.org/10.1143/JPSJ.80.072001

Lucas RE (1976) Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46. https://doi.org/10.1016/S0167-2231(76)80003-6

Malone TW and Bernstein MS (2015) *Handbook of collective intelligence*, MIT press.

Muth JF (1961) Rational expectations and the theory of price movements. *Econometrica* 29: 315–335. https://doi.org/10.2307/1909635

Nirei M and Aoki S (2016) Pareto distribution of income in neoclassical growth models. *Rev Econ Dyn* 20: 25–42. https://doi.org/10.1016/j.red.2015.11.002

Samuels JM (1965) Size and the growth of firms. *Rev Econ Stud* 32: 105–112. https://doi.org/10.2307/2296055

Sano K (2022) A binary decision model and fat tails in financial market. *Appl Sci* 12: 7019. https://doi.org/10.3390/app12147019

Tversky A and Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science* 185: 1124–1131. https://doi.org/10.1126/science.185.4157.112