# Power laws in market capitalization during the Dot-com and Shanghai bubble periods

Takayuki Mizuno

Takaaki Ohnishi

Tsutomu Watanabe

# Power laws in market capitalization during the Dot-com and Shanghai bubble periods

**Takayuki Mizuno, Takaaki Ohnishi,
Tsutomu Watanabe**

**Abstract** The distributions of market capitalization across stocks listed in the NASDAQ and Shanghai stock exchanges have power law tails. The power law exponents associated with these distributions fluctuate around one, but show a substantial decline during the dot-com bubble in 1997-2000 and the Shanghai bubble in 2007. In this paper, we show that the observed decline in the power law exponents is closely related to the deviation of the market values of stocks from their fundamental values. Specifically, we regress market capitalization of individual stocks on financial variables, such as sales, profits, and asset sizes, using the entire sample period (1990 to 2015) in order to identify variables with substantial contributions to fluctuations in fundamentals. Based on the regression results for stocks in listed in the NASDAQ, we argue that the fundamental value of a company is well captured by the value of its net asset, therefore a price book-value ratio (PBR) is a good measure of the deviation from fundamentals. We show that the PBR distribution across

T. Mizuno
National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan.
School of Multidisciplinary Sciences, SOKENDAI (The Graduate University for Advanced Studies), 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan.
PRESTO, Japan Science and Technology Agency, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan.
The Canon Institute for Global Studies, 1-5-1 Marunouchi, Chiyoda-ku, Tokyo, Japan.
E-mail: mizuno@nii.ac.jp

T. Ohnishi
Graduate School of Information Science and Technology, The University of Tokyo. Hongo, Bunkyo-ku, 113-8656 Tokyo, Japan.
The Canon Institute for Global Studies, 1-5-1 Marunouchi, Chiyoda-ku, Tokyo, Japan.

T. Watanabe
Graduate School of Economics, The University of Tokyo. Hongo, Bunkyo-ku, 113-0033 Tokyo, Japan.
The Canon Institute for Global Studies, 1-5-1 Marunouchi, Chiyoda-ku, Tokyo, Japan.

stocks listed in the NASDAQ has a much heavier upper tail in 1997 than in the other years, suggesting that stock prices deviate from fundamentals for a limited number of stocks constituting the tail part of the PBR distribution. However, we fail to obtain a similar result for Shanghai stocks.

# 1 Introduction

Since B. Mandelbrot identified the fractal structure of price fluctuations in asset markets in 1963 [1], statistical physicists have been investigating the economic mechanism through which a fractal structure emerges. Power laws is an important characteristic in the fractal structure. For example, some studies found that the size distribution of asset price fluctuations follows power law [2,3]. Also, it is shown that firm size distribution (e.g., the distribution of sales across firms) also follows power law [4–8]. The power law exponent associated with firm size distributions is close to one over the last 30 years in many countries [9,10]. The situation in which the exponent is equal to one is special in that it is the critical point between the oligopolistic phase and the pseudo-equal phase [11]. If the power law exponent less than one, the finite number of top firms occupy a dominant share in the market even if there are infinite number of firms.

One of the most important issues along this line of research is power laws associated with asset price fluctuations, and several models describing asset price dynamics were proposed [12,13]. In particular, asset price bubbles were regarded as an important research topic. For example, the PACK and LPPL models simulated the price fluctuations of a stock during bubble periods [14, 15]. In economics and finance, asset price bubbles are defined as the deviation of the price of an asset from its fundamental value. However, it is not easy to obtain information about fundamentals. For example, it is often stated by economists that the fundamental stock value of a company equals to the present discounted value of dividends delivered by the company in the years to come. However, it is hard to get a reliable estimate of future dividends, therefore it is next to impossible to estimate fundamental stock prices. Without accurate information on fundamentals, it is impossible to detect bubbles. This is a serious issue for policy makers, like governments and central banks, since the emergence and the burst of bubbles often lead to intolerable economic disasters, such as financial crisis.

The purpose of this paper is to propose a method to detect stock price bubbles in a timely manner. Our basic idea is closely related to a method we proposed as a way to detect bubbles in the context of real estate prices [16]. In this study on real estate prices, we looked for houses that are similar in various respects, including the location of a house, the size of a house, and the age

of a house. We argued that houses with similar attributes can be regarded as having similar fundamental values, therefore the prices for these houses should be similar if there is no bubbles in the housing market. Based on this idea, we looked at the distribution of house prices for houses with similar attributes, showing that it is close to a lognormal distribution during normal periods, but it has a heavy upper tail during bubble periods. In the present paper, we apply this idea to stock markets to detect stock price bubbles.

In this paper, we use a dataset compiled by the Thomson Reuters Corporation that covers daily market capitalization and annual income statements of all the listed firms in NASDAQ and SSE from 1990 to 2015. This period includes the 2000 dot-com and 2007 Shanghai bubbles. We focus on the distribution of market capitalization in NASDAQ and the Shanghai stock exchange (SSE) for these bubbles. Kaizoji et al. showed that the upper tail of stock price distribution in the Tokyo stock exchange grew fat during the dot-com bubble period [17, 18]. However, if we accurately investigate the firm size in stock markets, not only the price but also the outstanding shares must be taken into consideration because share consolidation and splitting often occurs in the market.

The rest of the paper is organized as follows. Section 2 examines the power law of market capitalization distribution using an expansion of the Castillo and Puig test [10, 19–21]. In Section 3 we observe that the power law index fluctuates around one, depending on economic conditions, and tends to become smaller during bubble periods. In Section 4, we find that net assets are most reflected in market capitalization for firms listed in NASDAQ during non-bubble periods. The price book-value ratio (PBR) distribution, which is defined as market capitalization divided by net assets, got fat during the bubble periods. These results suggest that speculative money is excessively concentrated on specific stocks during bubble periods. Section 5 concludes the paper.

## 2 Distribution of market capitalization

First, we show the cumulative distribution of market capitalization for the listed firms in a stock market. The distribution on December 31, 1997 in NASDAQ is shown in log-log plots in Fig. 1. The distribution is on a straight line that indicates a power law function. However, in this case, the distribution deviates from a straight line and is close to a log-normal distribution when the market capitalization is smaller than about $2.7 \times 10^8$ dollars:

$$P_>(x) \propto x^{-\mu} \text{ for } x \geq x_0, \tag{1}$$

where $x$ is the market capitalization, $\mu$ is a power law index, and $x_0$ is a threshold. The index estimated with the maximum likelihood method is $\mu = 1.0$ (Fig. 1). Such a power law with $\mu = 1.0$ is called Zipf's law.

Malevergne et al. [19], who expanded the likelihood ratio test between exponential distribution and the truncated normal distribution introduced by
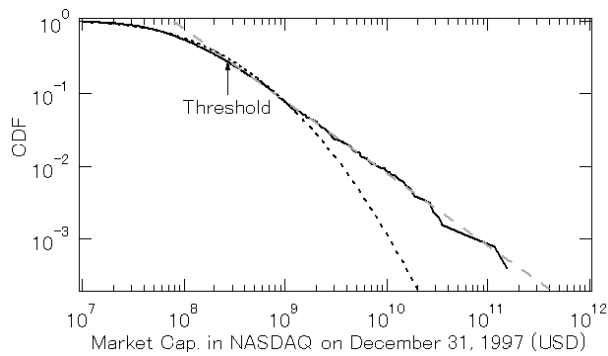
**Fig. 1 Cumulative distribution of market capitalization for listed firms in NAS-DAQ on December 31, 1997.** Dashed straight line and curve respectively express a power law distribution as $P_>(x) \propto x^{-1}$ and log-normal distribution with standard deviation of market capitalizations. Arrow indicates $x_0 = 2.7 \times 10^8$ dollars in Eq. (1).

Castillo and Puig [20], tested the null hypothesis where, beyond a threshold, a distribution's upper tail is characterized by a power law distribution against the alternative where the upper tail follows a log-normal distribution beyond the same threshold. This is known as the uniformly most powerful unbiased test. They identified the upper tail that follows a log-normal distribution or a power law distribution to detect a threshold by conducting this test.

The dashed line in Fig. 1 displays the threshold between a power law distribution and a log-normal distribution when the $p$-value of the significance level is set as 0.1 for Malevergne's test. Therefore, the upper tail of market capitalization distribution can be well approximated by the power law function. On the other hand, a small range of market capitalization follows the log-normal distribution, as shown in the dashed curve in Fig. 1.

## 3 Power law index of market capitalization distribution

The upper tail of market capitalization distribution gets fat if the speculative money is concentrated on specific stocks. Such concentration of money tends to occur during bubble periods. The power law index became smaller during the 2000 dot-com and 2007 Shanghai bubbles.

The means of market capitalization for all the listed firms in NASDAQ and SSE were respectively about $1.94 \times 10^9$ dollars on March 9, 2000 and $3.92 \times 10^9$ dollars on December 4, 2007 during their bubble periods. On the other hand, the means were respectively about $1.44 \times 10^9$ and $3.04 \times 10^9$ dollars on March 14, 2011 on the non-bubble periods. The means increased during the bubble periods.

Why does the mean increase during bubble periods? One possibility is that only a few firms increased the market capitalization drastically and raised the whole mean of market capitalization for all the listed firms. The black
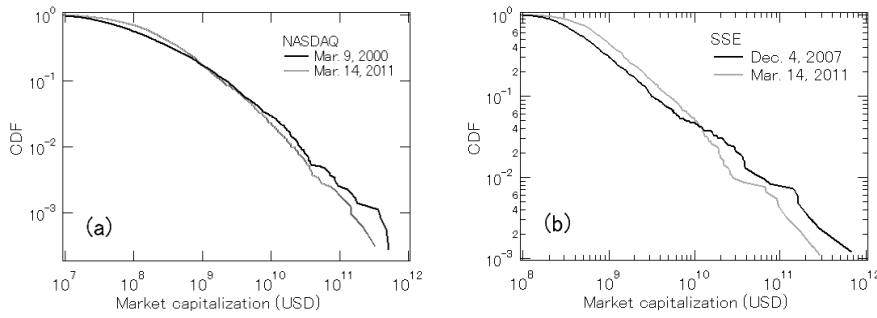
**Fig. 2** Cumulative distributions of market capitalization for listed firms: (a) in NASDAQ on March 9, 2000 and on March 14, 2011 and (b) in SSE on Decenber 4, 2007 and on March 14, 2011.

lines in Fig. 2 respectively show the market capitalization distributions for firms listed on the NASDAQ during the 2000 dot-com bubble and on the SSE during the 2007 Shanghai bubble. The gray lines express the distributions on March 14, 2011 during non-bubble periods. The cumulative probability on the vertical axis at which the black line intersects with the gray line is $P_>(x = 2 \times 10^9) \approx 0.1$ in the dot-com bubble case. The top 10% market capitalization in 2000 was higher than that in 2011, although the bottom 90% market capitalization in 2000 was lower than that in 2011. Such a characteristic was also observed in the Shanghai bubble case (Fig. 2(b)). The cumulative probability of its crossing point is $P_>(x = 1.2 \times 10^{10}) \approx 0.04$. These results suggest that speculative funds concentrate on a very small set of stocks, leading to stock price bubbles.

The concentration of speculative money changes the distribution slope. Fig. 3 shows the NASDAQ composite index, the SSE composite index, and the time series of the power law index in NASDAQ and SSE. The power law index, which fluctuated around one depending on the economic conditions, fell significantly when the increase of each composite index started during the dot-com and 2007 Shanghai bubble periods and remained small until each composite index steadied after the bubble burst.

## 4 Fundamental-adjusted market capitalization in NASDAQ

In economics, a financial bubble is defined by the gap between market and fundamental prices. We show that the gap expanded during the 2000 dot-com bubble period.

First, we look for a firm's fundamentals that are mainly reflected in its market capitalization during non-bubble periods. We chose the following key financial variables, total assets, net assets, total revenue, operating income, net income, operating cash flow, and number of employees, as candidates of firm fundamentals and investigated the correlation between market capitalization
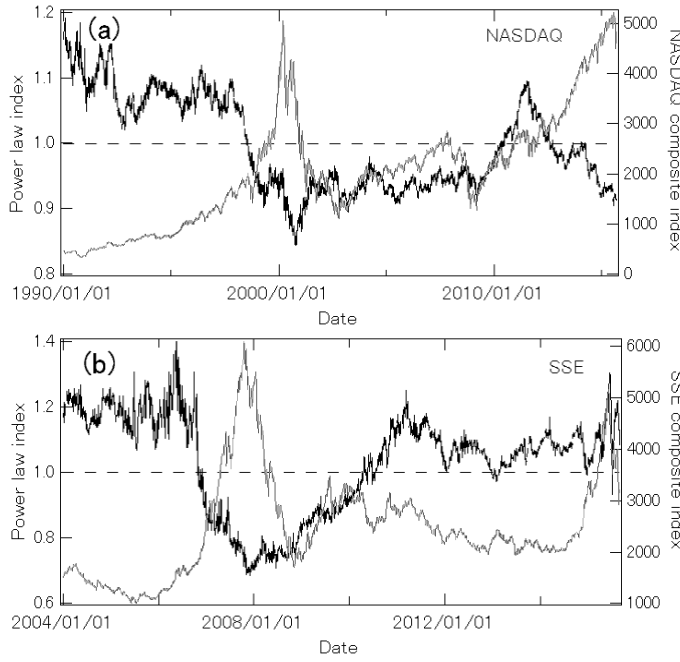
**Fig. 3 (a) Power law index in NASDAQ and NASDAQ composite index and (b) Power law index in SSE and SSE composite index.** Black and gray lines respectively express power law and composite indexes. Dashed lines indicate the power law index is one.

and each key financial variable during non-bubble periods. Table 1 expresses the Kendall and Pearson correlation coefficients for firms listed on NASDAQ in 1997 and 2004. The correlation coefficients between market capitalization and net assets are the largest. However, other correlation coefficients are also high, suggesting the possibility of spurious correlation. To cope with this problem, we conduct random forest regression with market capitalization of individual firms as dependent variable and other financial variables as independent variables.

We provide the regression coefficients of random forests, which are hardly affected by multicollinearity even though we used the data that caused it. Random forests can be used to rank the importance of the explanatory variables in a regression by estimating the explained variable using regression trees without part of the explanatory variables. We set market capitalization and the financial variables to the explained variable and the explanatory variables. Fig. 4(a) displays the time series of the importance of the financial variables from 1995 to 2013 in NASDAQ. Net assets are the most important in all years except 2011. Other financial variables are not important for most years. On the other hand, in SSE, both net assets and net income are critical. Fig. 4(b) also expresses the importance of financial variables from 2000 to 2013 in SSE. These results suggest that it is highly likely that the high correlations observed

**Table 1 Kendall and Pearson correlation coefficients between market capitalization and each key financial variable in 1997 and 2004 in NASDAQ.** Pearson correlation coefficients are calculated using logarithmic values of market capitalization and financial variables without negative values.

| | 1997 | | 2004 | |
| --- | --- | --- | --- | --- |
| Financial variable | Kendall | Pearson ($> 0$) | Kendall | Pearson ($> 0$) |
| Total assets | 0.461 | 0.684 | 0.526 | 0.755 |
| Net assets | 0.622 | 0.824 | 0.681 | 0.880 |
| Total revenue | 0.457 | 0.632 | 0.523 | 0.680 |
| Operating income | 0.473 | 0.820 | 0.452 | 0.841 |
| Net income | 0.410 | 0.805 | 0.420 | 0.851 |
| Operating cash flow | 0.351 | 0.745 | 0.449 | 0.829 |
| Number of employees | 0.396 | 0.582 | 0.466 | 0.658 |

between market capitalization and other financial variables may be spurious ones.

Next, we introduce fundamental-adjusted market capitalization. In NASDAQ, since only net assets are the main fundamental that is reflected in market capitalization, we introduce market capitalization adjusted by net assets. In finance, net asset-adjusted market capitalization is called the PBR:

$$PBR_i(t) = \frac{x_i(t)}{A_i(t)}, \tag{2}$$

where $x_i(t)$ is the market capitalization of firm $i$ on the settlement day in year $t$ and $A_i$ is its net assets in year $t$. Fig. 5 shows the distributions of $PBR(1997)$ in the pre-bubble period, of $PBR(1997)$ in the bubble period, and of $PBR(2004)$ in the post-bubble period in NASDAQ. The distribution became fat during the bubble period and returned to its former position after the bubble burst.

Next, we focus on SSE. Unlike NASDAQ (Fig. 4(b)), its market capitalizations reflect two financial variables: net assets and net income. Because net income is rarely around zero, the market capitalization divided by the net income becomes extremely big at that time. The upper tail of the distribution of divided market capitalization sensitively responds to the fluctuation of net income. Future work will propose market capitalization adjusted by both net assets and net income.

## 5 Conclusion

We showed the distributions of market capitalization in NASDAQ and SSE. The upper tails of the distributions follow a power law. The power law index, which fluctuates around one depending on the economic conditions, became small during the 2000 dot-com and 2007 Shanghai bubble periods, suggesting that speculative money was excessively concentrated on a very small set of stocks, leading to stock price bubbles.
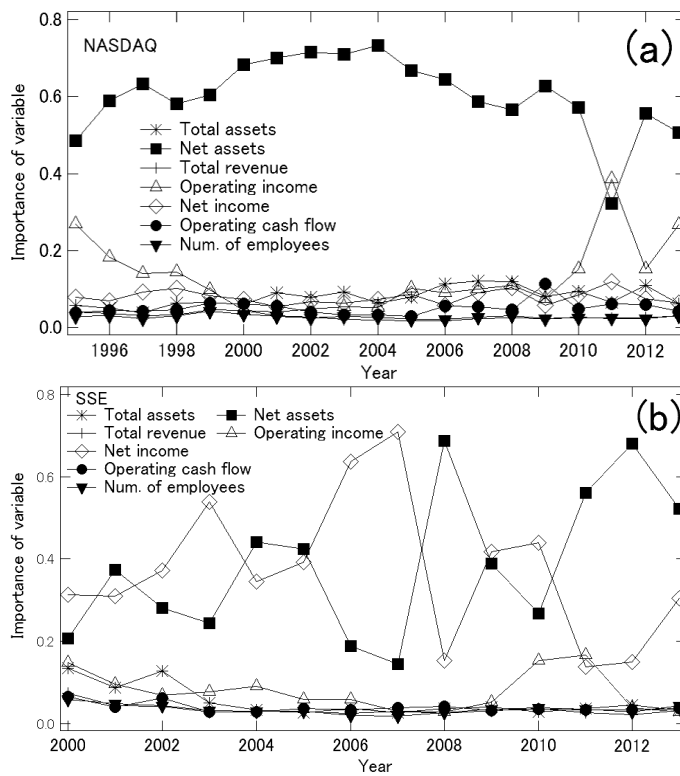
**Fig. 4 Importance of financial variables: (a) from 1995 to 2013 in NASDAQ and (b) from 2000 to 2013 in SSE.** $*$, $\blacksquare$, $+$, $\triangle$, $\diamondsuit$, $\bullet$, and $\blacktriangledown$ are respectively total assets, net assets, total revenue, operating income, net income, operating cash flow, and number of employees.



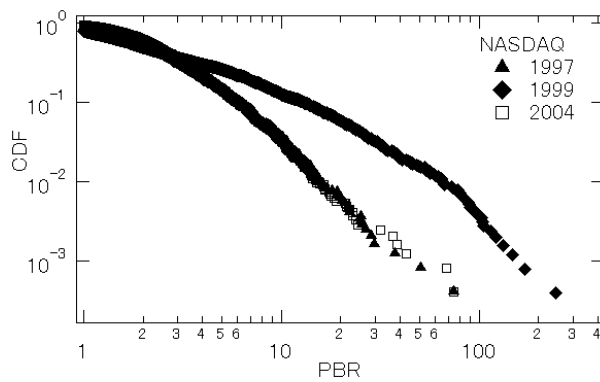**Fig. 5 Cumulative distributions of PBR in NASDAQ in ($\blacktriangle$) 1997, ($\blacklozenge$) 1999, and ($\square$) 2004.**

In economics and finance, a stock price bubble is defined by the gap between firm sizes in the stock market and in real economies. We used market capitalization and financial variables to estimate the firm sizes in stock markets and real economies. By using the regression coefficient of random forests for market capitalization and financial variables, we found that net assets are most reflected in the market capitalization for NASDAQ firms. For such firms, PBR is defined as market capitalization divided by net assets. The PBR distribution also got fat during the dot-com bubble period. This result means that the gap between firm sizes in asset markets and in real economies widened during the bubble period. This may be a useful tool for policy makers like governments and central banks to detect stock price bubbles. Note that changes in the PBR distribution can be monitored at the daily, or even higher frequencies, so policy makers will be able to evaluate the risk of asset price bubbles almost at a real time basis.

Both net assets and net income are greatly reflected in market capitalization for SSE firms. Market capitalization, divided by net income, becomes extremely big when the net income is close to zero. Therefore, the upper tail of the distribution of divided market capitalization sensitively responds to the fluctuation of net income. One future work will propose market capitalization that is adjusted by net assets and net income to investigate the 2007 Shanghai bubble.

Although market capitalization is made public every day, financial variables are usually announced only quarterly and annually. This difference in time scale complicates estimating daily gaps between firm sizes in stock markets and real economies. Another future work will nowcast the key financial variables every day.

# References

1. Mandelbrot B (1963) The Variation of Certain Speculative Prices. The Journal of Business 36(4):394-419.
2. Mantegna RN & Stanley HE (2000) Introduction to Econophysics: Correlations & Complexity in Finance. Cambridge University Press, Cambridge.
3. Mizuno T, Kurihara S, Takayasu M & Takayasu H (2003) Analysis of high-resolution foreign exchange data of USD-JPY for 13 years. Physica A 324:296-302.
4. Stanley MHR, Buldyrev SV, Havlin S, Mantegna R, Salinger MA & Stanley HE (1995) Zipf plots and the size distribution of Firms. Economics Lett. 49:453-457.
5. Axtell RL (2001) Zipf Distribution of U.S. Firm Sizes. Science 293:1818-1820.
6. Okuyama K, Takayasu M & Takayasu H (1999) Zipf's Law in Income Distribution of Companies. Physica A 269:125-131.
7. Mizuno T, Ishikawa A, Fujimoto S & Watanabe T (2012) Power Laws in Firm Productivity. Progress of Theoretical Physics Supplement 194:122-134.
8. Clementi F & Gallegati M (2016) The Distribution of Income and Wealth: Parametric Modeling with the k-Generalized Family. Springer.

9. Mizuno T, Katori M, Takayasu H & Takayasu M (2002) Statistical and Stochastic Laws in the Income of Japanese Companies. Empirical Science of Financial Fluctuations -The Advent of Econophysics. Springer Verlag, Tokyo:321-330.
10. Fujimoto S, Ishikawa A, Mizuno T & Watanabe T (2011) A New Method for Measuring Tail Exponents of Firm Size Distributions. Economics 5:2011-19.
11. Aoyama H, et al. (2010) Econophysics and Companies: Statistical Life and Death in Complex Business Networks. Cambridge University Press: 61-63.
12. Richmond P, Mimkes J & Hutzler D (2013) Econophysics and Physical Economics. Oxford University Press.
13. Abergel F, et al. (2016) Limit Order Books. Cambridge University Press.
14. Sornette D (2004) Critical phenomena in natural sciences: chaos, fractals, self organization, and disorder: concepts and tools. Springer.
15. Takayasu M, Watanabe K, Mizuno T & Takayasu H (2010) Theoretical Base of the PUCK-Model with application to Foreign Exchange Markets. Econophysics Approaches to Large-Scale Business Data and Financial Crisis. Springer:79-100.
16. Ohnishi T, Mizuno T, Shimizu C & Watanabe T (2012) Power laws in real estate prices during bubble periods. International Journal of Modern Physics:Conference Series 16:61-81.
17. Kaizoji T (2006) Power laws and market crashes: Empirical laws on bursting bubbles. Progress of Theoretical Physics Supplement 162:165-172.
18. Kaizoji T (2006) Precursors of Market Crashes: Empirical laws of the Japan's internet bubbles. European Physical Journal B 50: 123-127.
19. Malevergne Y, Pisarenko V & Sornette D (2011) Testing the Pareto against the Lognormal Distributions with the Uniformly Most Powerful Unbiased Test Applied to the Distribution of Cities. Physical Review E 83:036111.
20. Del Castillo J & Puig P (1999) The Best Test of Exponentiality against Singly Truncated Normal Alternatives. Journal of the American Statistical Association 94(446):529-532.
21. Hisano R & Mizuno T (2010) Sales Distribution of Consumer Electronics. Physica A 390:309-318.