Dynamics of US Buffett Indicator *

Hiro Asano

December 1, 2024

Points of this study

1 TMC and GDP are both nonstationary, as they are known.

2 The Buffett indicator, more precisely $lnS - \beta\, lnY$, is cointegrated with $\beta > 1$.

3 The Buffett indicator could be mean reverting.

This study investigates the dynamic relationship of two markets: the goods market and the stock market. It uses Gross Domestic Product (GDP), and total market capitalization (TMC) as size measures for those markets. GDP represents the value of all goods and services domestically produced, and TMC represents the aggregate market capitalization of all publicly traded firms. Both TMC and GDP exhibit dynamics of long-term trends and short-term fluctuations.

The study assumes that a relationship between TMC and GDP should exist. Value added is sales revenue less the costs of intermediate goods such as materials and energy. The profit is value added, less labor costs and depreciation. These items are proportional to output to some extent. In particular, value-added and profits are roughly proportional. Investors evaluate a firm based on its profit and holding assets, and contemplate its stock price. Thus, a firm's profit should reflect on its market capitalization, which is the product of its stock price and the number of outstanding stocks.

GDP is the aggregate value-added of all firms operating in a country. Profits are a part of the value added. If the discount rate and annual profits are constant over the years, a stock price theoretically equals annual profits divided by the discount rate. In other words, the discount rate's inverse is the theoretical value of the price-earnings ratio (P/E ratio). TMC is the aggregate of all firms' market capitalizations. In this way, TMC and GDP are related to each other.

The study analyzes the dynamic relationship between TMC and GDP by time series econometrics. TMC and GDP are both known as nonstationary or unit-root processes. When nonstationary time series data is regressed on another nonstationary time series, the regression can be spurious or nonsense. On the other hand, when both are related, nonstationary time series data could be cointegrated, and regression with them is possibly super-consistent. This study also investigates whether the indicator

---

* This is a preliminary and incomplete draft. The author is still working on many issues and any suggestions are welcome. (Email: asanohiro559@gmail.com)

is the mean-reverting process or whether the Buffett indicator shows the regression to the mean. If the indicator is the mean-reverting process, it tends to return to the long-run equilibrium level.

1 Data

1.1 Wilshire 5000 as Total Market Capitalization and Gross Domestic Product

This study analyses the US economy's total market capitalization (TMC) and Gross Domestic Product (GDP). The study employs the Wilshire 5000 index for TMC. The former measures the US stock markets, while the latter measures the US economy. The mathematical symbols of both data are, respectively, S and Y, and both are natural-logged. The upper panel of Figure 1 shows both time-series data. TMC and GDP data have been increasing in the long term, while TMC has fluctuated more vigorously than GDP in the short term. Also, it is shown that TMC has risen faster than GDP.

1.2 Buffett Indicator (BI)

The Buffett indicator is defined as the ratio of TMC to GDP. [1)]

$$BI \equiv \frac{S}{Y} = exp(lnS - lnY) \tag{1}$$

Its value is supposed to equal unity. When the stock price index is either high or low, the Buffett indicator is also high or low. The lower panel of Figure 1 shows the Buffett indicator. The indicator has also increased, resulting from TMC rising faster than GDP. This study focuses on the relation between lnS and lnY. The relation of the Buffett Indicator is $lnS - lnY$. However, to incorporate the upward trend of the Buffett Indicator, the study modifies the indicator to $lnS - \beta\, lnY$ with $\beta > 1$.

2 Nonstationary Time Series Econometrics

Both TMC and GDP are known to be nonstationary or to have unit roots. When one nonstationary time series data is regressed on another nonstationary time series, the regression could be spurious, i.e., it could show seemingly good results even though the data are irrelevant.

2.1 Nonstationarity

This section investigates nonstationarity of both data using the Box-Jenkins methodology and the augmented Dickey-Fuller test as a unit-root test.

1) Buffett and Loomis, 2001

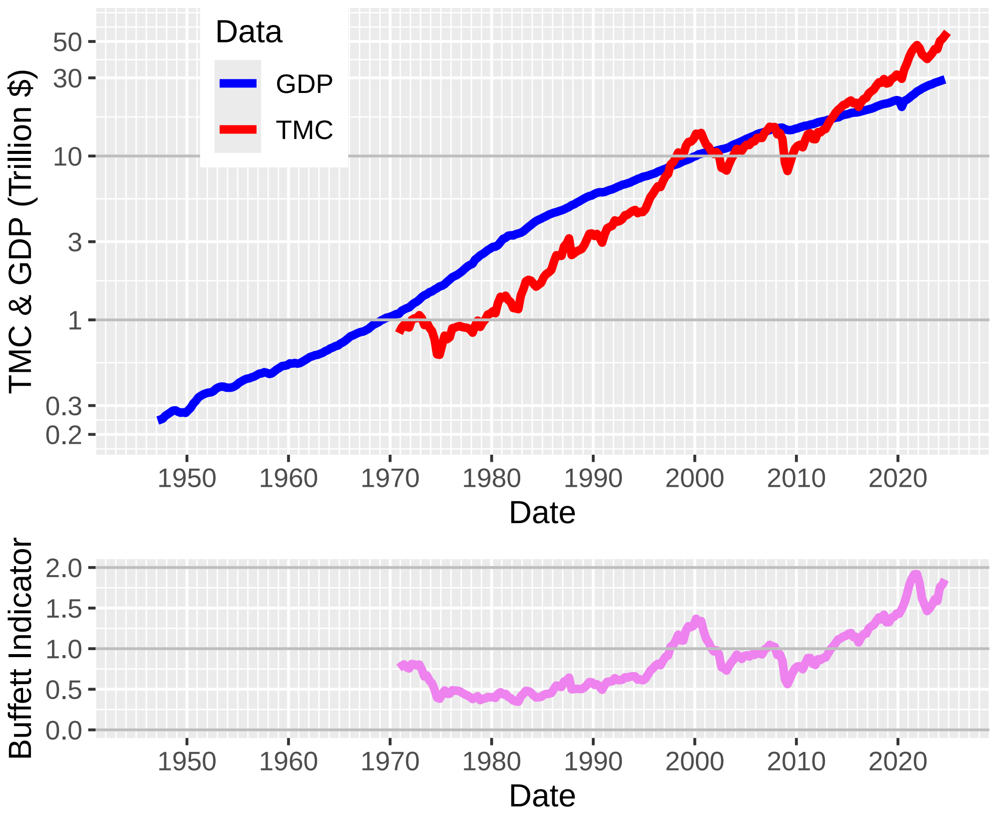


Figure1: Total market capitalization, GDP, and Buffett Indicator

2.1.1 Box-Jenkins Methodology

The Box-Jenkins methodology is a four-step procedure.

Step 1 Transform data to become data stationary, often taking differences.

Step 2 Determine the order of an autoregressive moving-average (ARMA) model.

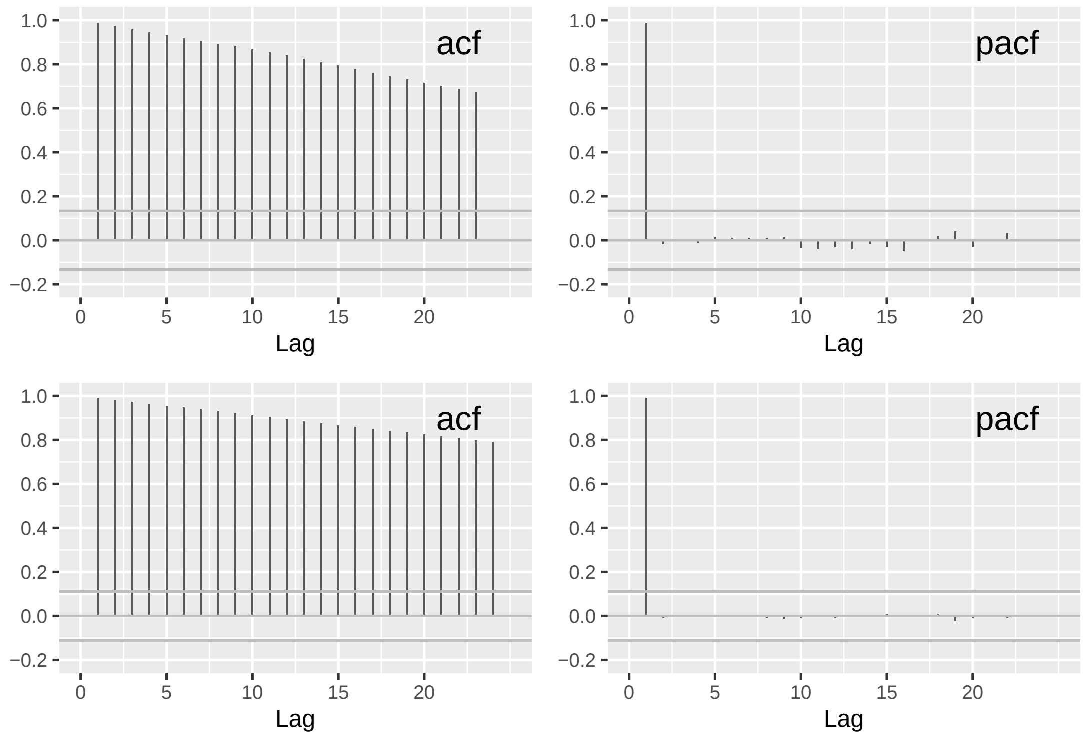
Step 3 Estimate the parameters of the ARMA process.

Step 4 Perform diagnostic analysis.

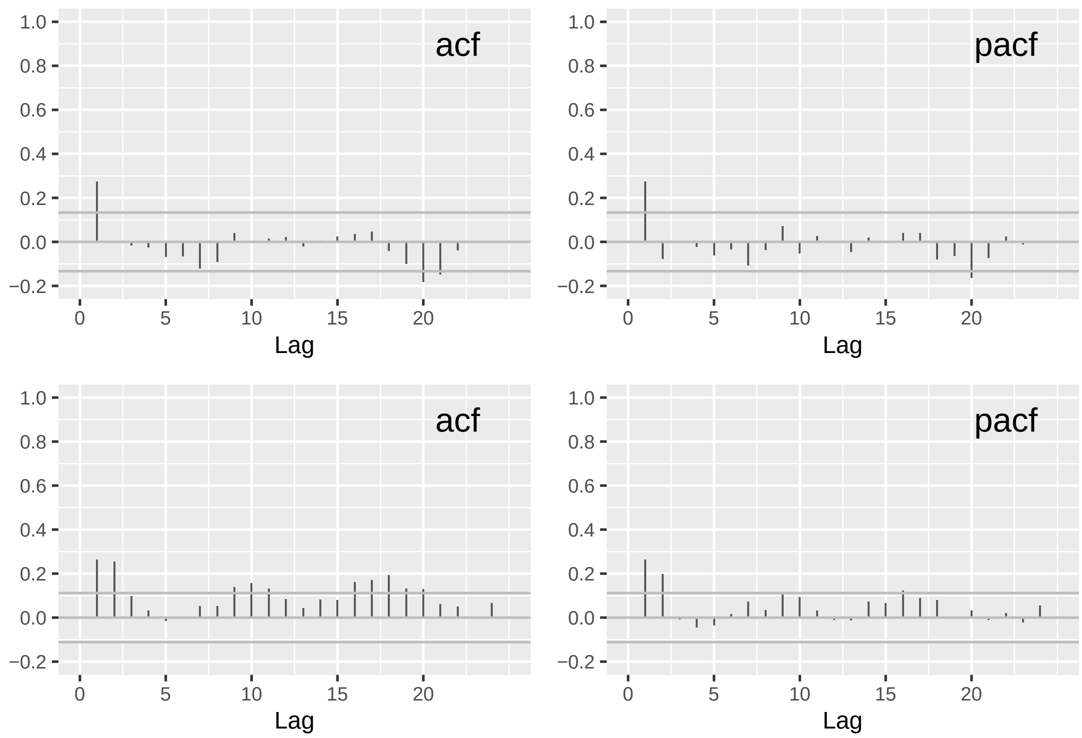
Table1 shows patterns of the autocorrelation function (acf) and the partial autocorrelation function (pacf) for AR(p), MA(q), and ARMA processes. The upper panel of Figure2 shows the correlograms of lnS_t and lnY_t. Both data show persisting acf, while their pacf's are insignificant. They are likely AR(1) processes, and possibly unit root processes. However, after both data are first-differenced, their acf's also become insignificant, as the lower panel of Figure2 shows.

2.1.2 Unit-root tests

Because both lnS_t and lnY_t seem nonstationary, this subsection conducts the augmented Dickey-Fuller (ADF) test for stationarity. Table2 shows its results. The ADF test concludes that lnS_t and lnY_t reject the null hypothesis of nonstationarity at the 5% significant level, while ΔlnS_t and



(a) Correlograms of lnS_t (above) and lnY_t (below)



(b) Correlograms of ΔlnS_t (above) and ΔlnY_t (below)

Figure2: Correlograms of lnS_t and lnY_t

Table1: Correlation patterns of ARMA process

Process	acf	pacf
AR(p)	Infinite: damping out	Finite: cuts off after lag p
MA(q)	Finite: cuts off after lag q	Infinite: damping out
ARMA	Infinite: damping out	Infinite: damping out

Table2: Augmented Dickey-Fuller test

Variable	Test statistic	p value	H_0	T
lnS_t	-2.6939698	0.2848389	nonstationary	217
ΔlnS_t	-6.2456496	< 0.01	nonstationary	216
lnY_t	0.0305701	> 0.99	nonstationary	311
ΔlnY_t	-5.9410323	< 0.01	nonstationary	310

ΔlnY_t fail to reject the hypothesis at the 5% significance level. In other words, lnS_t and lnY_t are nonstationary or unit root processes, while ΔlnS_t and ΔlnY_t are stationary.

2.2 Linear Regressions and Cointegration Analyses

The research subject of this study is the Buffett indicator, i.e., $lnS_t - lnY_t$ or, more specifically, $lnS_t - \beta\ lnY_t$. This section conducts linear regression analyses.

2.2.1 Simple Linear Regression Model

First, the study investigates the simple linear regression model (SLRM), and its regression equation is the following:

$$lnS_t = \alpha_{SLRM} + \beta_{SLRM}\ lnY_t + u_t \tag{2}$$

As table3 shows, the results of the SLRM estimation seem good. Namely, the coefficient estimate of β_{SLRM}, is significant, and R^2 is high and close to unity. However, the Durbin-Watson statistic is less than R^2, which is a good rule of thumb to suspect that the regression is spurious.

Table3: Simple linear regression model

Coefficient / test	Estimate / test statistic	p value	H_0
α_{SLRM}	-3.7336868	1.6807021E-49	$= 0$
β_{SLRM}	1.3893131	3.8439479E-143	$= 0$
F	4249.5967084	3.8439479E-143	all coefficients equal zero
Durbin-Watson	0.0535566	1.4498556E-47	zero autocorrelation

T = 216; R^2 = 0.9520566; $\overline{R^2}$ = 0.9518326

2.2.2 Error-correction Model

This section employs the error correction model (ECM) to improve the SLRM regression. Its regression equation is the following:

$$\Delta lnS_t = \alpha_{ECM} + \beta_{ECM}\ \Delta lnY_t + \gamma_s\ lnS_{t-1} + \gamma_y\ lnY_{t-1} + v_t \tag{3}$$

Table4 shows the results of the ECM regression. The estimates of the coefficients, β_{ECM}, γ_s, and γ_y are siginificant, and R^2 is less than the Durbin-Watson statistic. Also, the ratio $-\gamma_y\ /\ \gamma_s$ ($\approx$ 1.6666) is close to the estimates of coefficients, β_{ECM} and β_{SLRM}. By rewriting the ECM regression equation, we have the following:

$$(\Delta lnS_t - \beta_{ECM}\ \Delta lnY_t) = \alpha_{ECM} + \gamma_s \left[lnS_{t-1} - \left(-\frac{\gamma_y}{\gamma_s} \right) lnY_{t-1} \right] + v_t \tag{4}$$

The study compares the equilibrium error term on the right-hand side and the transient error term on the left-hand side by the delta method. The testing null hypothesis is $\beta_{ECM} = -\gamma_y/\gamma_s$, or $\beta_{ECM} + \gamma_y/\gamma_s = 0$. Table5 shows its result. The hypothesis test by the delta method fails to reject the hypothesis, or in other words, these terms are statistically equal.

Table4: Error correction model

Coefficient / test	Estimate / test statistic	p value	H_0
α_{ECM}	-0.2542026	0.001588	$= 0$
β_{ECM}	1.5070731	8.5912221E-5	$= 0$
γ_s	-0.0403735	0.011012	$= 0$
γ_y	0.0672852	0.0034539	$= 0$
F	6.8034087	2.1257972E-4	all coefficients equal zero
Durbin-Watson	1.5018376	6.9368337E-5	zero autocorrelation

$T = 215$; $R^2 = 0.0881993$; $\overline{R^2} = 0.0752353$

Table5: Comparison of equilibrium and transition terms

Relation	Test statistic	p value	H_0
$\beta_{ECM} + \gamma_y/\gamma_s$	-0.1594959	0.663689	$= 0$

2.2.3 Stationarity of Residuals

When the regression is spurious, its residuals are nonstationary. Therefore, the study investigates the SLRM and the ECM residuals using the correlograms and the Engel-Granger ADF (EG-ADF) test. Figure3 shows the correlograms of the SLRM and the ECM residuals. The acf of the SLRM residuals, $\hat{u}$, damps out gradually, while all three other functions cut off after the first lag. In other words, residuals, $\hat{u}$ and $\hat{v}$, are both stationary and, because the SLRM and the ECM regressions both yield stationary residuals, lnS_t and lnY_t are cointegrated.

Table6 shows the results of the EG-ADF test. The EG-ADF test and the ADF test have similar procedures but different criteria. Table7 shows the criteria of the EG-ADF test. Both $\hat{v}$ and $\hat{u}$ reject the null hypothesis of nonstationarity, although the ECM residuals, $\hat{v}$, more strongly reject the hypothesis than the SLRM residuals, $\hat{u}$.

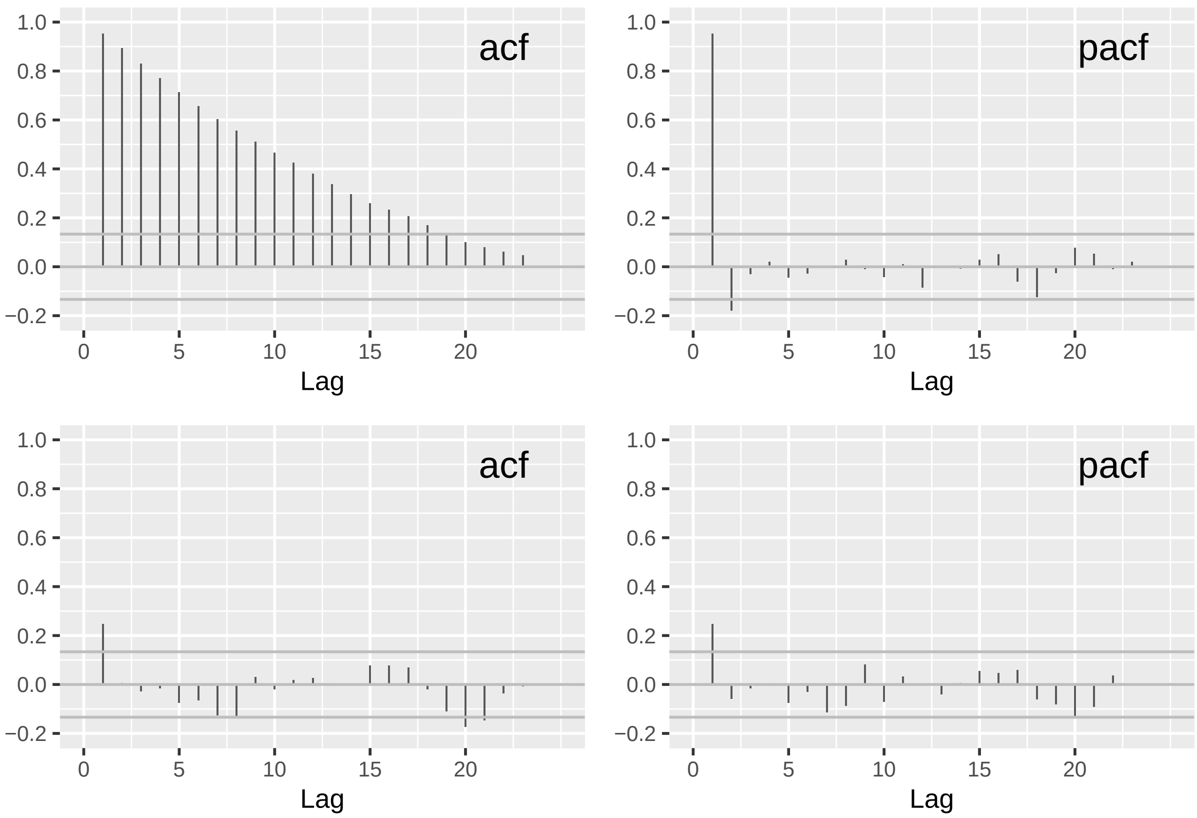


Figure3: Correlograms of SLRM residuals (above) and ECM residuals (below)

Table6: Engle-Granger augmented Dickey-Fuller tests

Residuals	Statistic	p value	H_0	T
$\hat{u}$	-3.6080775	(0.0336199)	nonstationary	216
$\hat{v}$	-6.4017524	($<$ 0.01)	nonstationary	215

Remark: the p values require adjustments.

Table7: Critical values for EG-ADF statistics

Number of explanatory variables	10%	5%	1%
1 (for $\hat{u}$)	-3.12	-3.41	-3.96
3 (for $\hat{v}$)	-3.84	-4.16	-4.73

2.2.4 Johansen Test

The Johansen test is for cointegration analysis, which also employs the ECM. It has two tests: the trace test and the maximum eigenvalue test. Table8 shows that both tests reject the null hypothesis of zero cointegration vector but fail to reject one or more cointegration vectors. In other words, the Johansen test concludes that lnS_t and lnY_t have one cointegrating vector.

Table8: Johansen tests

Test	Test statistic	10%	5%	1%	H_0
Trace	5.2100697	6.5	8.18	11.65	≥ 1
	44.1625426	15.66	17.95	23.52	$= 0$
Maximum eigenvalue	5.2100697	6.5	8.18	11.65	≥ 1
	38.952473	12.91	14.9	19.19	$= 0$

3 Buffett Indicator as Mean-reverting Process

If $\beta_{ECM} = \gamma_y/\gamma_s$, we can rewrite equation (4) as a following mean-reverting process with $\kappa \equiv -\gamma_s$, $\mu \equiv -\alpha_{ECM}/\gamma_s$, and $lnBI_t \equiv lnS_t - \beta_{ECM}\ lnY_t$;

$$\Delta lnBI_t = \kappa\left(\mu - lnBI_{t-1}\right) + v_t \tag{5}$$

The mean-reverting process is also called the Ornstein-Uhlenbeck process. Table9 shows the estimates of the coefficients, κ and μ. The estimate for coefficient κ is positive and statistically significant, so that, when $lnBI_t > \mu$, $\Delta lnBI_t$ likely becomes negative. In other words, when $lnBI_{t-1}$ is greater than its mean , μ, $lnBI_t$ is likely to less than $lnBI_{t-1}$, i.e., $lnBI_t$ is mean reverting.

If residuals, the ECM residuals, $\hat{v}_t$, or the SLRM residuals, $\hat{u}_t$, is normally distributed, v_t or u_t could be a Wiener process. As Figure4 shows, the ECM residual, $\hat{v}_t$ are skewed to the left, while the

Table9: Mean-reverting Process

Coefficient	Estimate	p value	H_0
κ	0.0403735	0.011012	$= 0$
μ	-6.2962747	2.0947327E-5	$= 0$

SLRM residuals, $\hat{u}_t$ are skewed to the right. Here, the ECM residuals, $\hat{v}_t$, seem to have a normal distribution with four outliers. As Table10 shows, the Shapiro-Wilk test for normality rejects the null hypothesis for both $\hat{v}_t$ and $\hat{u}_t$.

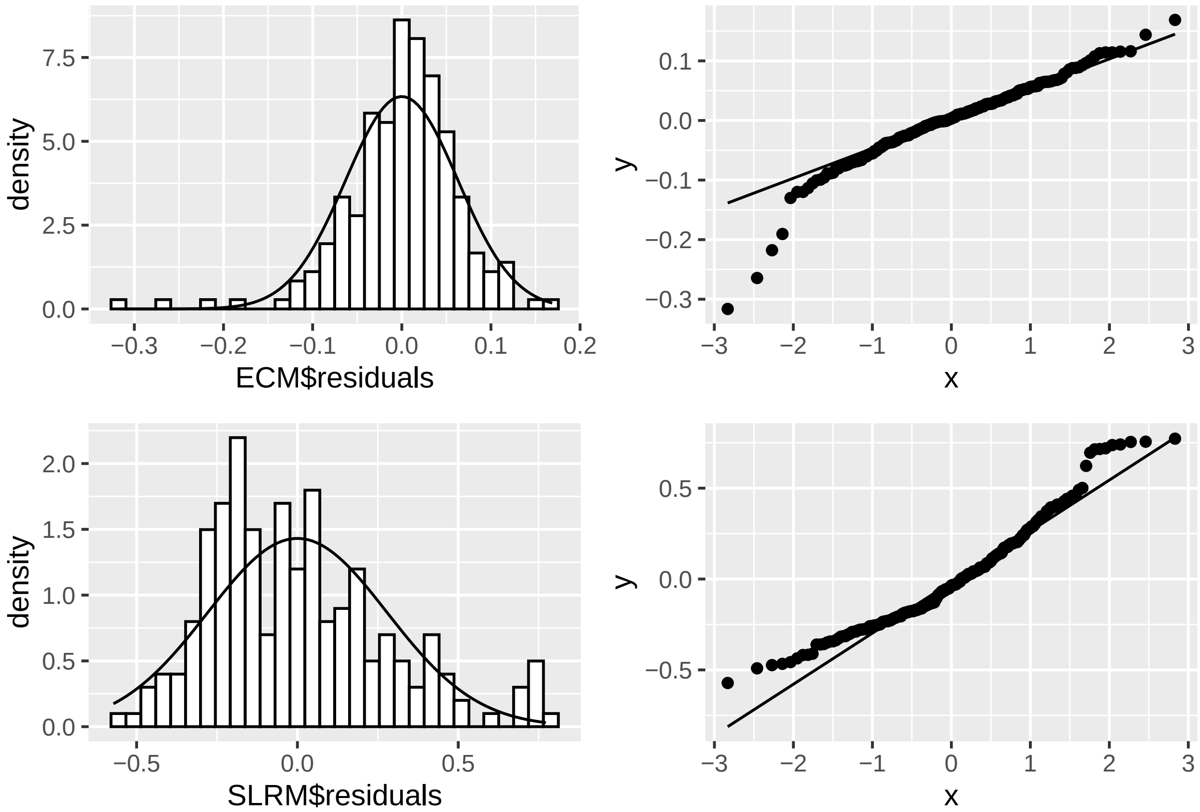


Figure4: Histograms and QQ-plots of $\hat{v}$ (above) and $\hat{u}$ (below)

Table10: Shapiro-Wilk tests for normality

Residuals	Test statistic	p value	H_0
$\hat{v}$	0.9391081	7.9818197E-8	normality
$\hat{u}$	0.9548824	2.5682555E-6	normality

4 Future Reserch

Possible future research would be three-hold: incorporating super-consistent standard errors into analysis, testing whether the Buffett indicator could be the Wiener process with the Poisson process, and testing the hypothesis that the indicator is mean reverting, and if so, what the mean is.

When time series data are cointegrated, a regression with such data could be super-consistent.[2)] The convergence speed of coefficient estimates is $1/T$ rather than $1/\sqrt{T}$. Then, hypothesis tests could yield different results. In particular, $\beta_{ECM} \neq -\gamma_y/\gamma_s$, i.e., the equilibrium error term and the transient error term are different, as opposed to this study.

The ECM residuals seem to have a normal distribution with four outliers. By dropping those outliers, its distribution could be normal. In addition, The four outliers may follow the Poisson process.

If $\beta_{ECM} \neq -\gamma_y/\gamma_s$, then equation (4) can be rewritten as follows:

$$\begin{aligned}(\Delta lnS_t - \beta_{ECM}\ \Delta lnY_t) &= \alpha_{ECM} + \gamma_s \left[lnS_{t-1} - \left(-\frac{\gamma_y}{\gamma_s} \right) lnY_{t-1} \right] + v_t \\ &= \gamma_s \left[\frac{\alpha}{\gamma_s} + \left(\beta_{ECM} + \frac{\gamma_y}{\gamma_s} \right) lnY_{t-1} + (lnS_{t-1} - \beta_{ECM}\ lnY_{t-1}) \right] + v_t\end{aligned}$$

Then,

$$\Delta lnBI_t = \kappa \left[\mu - \left(\beta_{ECM} + \frac{\gamma_y}{\gamma_s} \right) lnY_{t-1} - lnBI_{t-1} \right] + v_t \tag{6}$$

Therefore, $lnBI_t$ has its mean to revert, which reduces the bracket term of the right-hand side to zero. The reverting mean can be written as follows:

$$lnBI_t = \mu - \left(\beta_{ECM} + \frac{\gamma_y}{\gamma_s} \right) lnY_t \tag{7}$$

Thus, the Buffett indicator tends to increase along with nominal GDP.

5 References

- Buffett, Waren, and Carol Loomis, "Warren Buffett on the Stock Market," the *Fortune* Magazine, December 2001.
- Stock, James H., "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica*, September 1987.

2) Stock, 1987