M1, M2, and the U.S. Equity Exchanges

Ali M. Parhizgari¹
Duong Nguyen²

Abstract
We study the relative positions of M1 and M2 in light of their relationships with four U.S. equity exchanges: S&P500, Dow Jones Industrial, Nasdaq, and Wilshire 5000 composite. It is demonstrated that a long-term equilibrium relationship does indeed exist. Short-run dynamics are also considered and are found to be temporary departures from the long-run equilibrium. Based on a model, which yields robust estimated results and is thus considered well behaved, the direction of causality is established. The model is then put further to test to check the predictive power of the M1 and M2 money aggregates. Based on a set of in- and out-of-sample forecast experiments, the results overwhelmingly indicate that M2 is a better predictive measure and hence a superior indicator than M1. The policy implications of these findings in light of the post financial crisis and the November 2010 US Fed “quantitative easing” policies are discussed.

Keywords: Money supply, M1, M2, Stock return, Granger causality, Error Correction Model, Forecasting.

JEL Classification: E50, E51, E52.

¹ Corresponding author: Ali M. Parhizgari, Ph.D., Professor of Finance and International Business and Ingersoll Rand Professor, Finance Department, Florida International University, Miami, FL 33199, USA. Email: parhiz@fiu.edu; Phone: (305) 348-3326.
² Duong Nguyen, Ph.D., Assistant Professor of Finance, Department of Accounting and Finance, Charlton College of Business, University of Massachusetts Dartmouth, 285 Old Westport Road, North Dartmouth, MA 02747, USA. Email: dnguyen2@umassd.edu; Phone: 508-999-8070.

We thank professor Ephraim Clark (the editor) and an anonymous referee for their many helpful suggestions and comments while the paper was in preparation. All the remaining errors and omissions, however, are ours.
1 – Introduction

Money supply and stock returns have widely been studied. Initial research in this area was mostly preoccupied with simply establishing a relationship between these variables [see, for instance, Homa and Jaffee, (1971); Hamburger and Kochin (1972); Rozeff (1974); Gupta (1974); Wright (1976); and Kraft and Kraft (1977)]. In the 1980’s and early 1990’s, refinements in modeling and, in particular, the causality issue were predominant [Lynge (1981); Sorensen (1982); Pearce and Roley (1983); Canto, et al. (1985); Darrat (1987); Jones and Uri (1987); Hashemzadeh and Taylor (1988); Bailey (1989); Foote (1989); and Dhakal, Kandil and Sharma (1993)]. Since the mid 1990’s and until recently, stability, relevance, and the predictive power of variants of money measures and/or stock returns have been in focus [see, for example, Mehra (1993); Fletcher and Chen (1993); Campbell and Kyle (1993); Choudhry (1996); Siklos and Anusiewicz (1998); Thornton (1998); Fleming, Kirby and Ostdiek (1998); Duca (2000); Bennett and Sias (2001); Huang, and Shen (2002); and Bahmani-Oskooee and Chomsisengphet (2002)].

The early 1990’s also witnessed substantial ebbs in the use of monetary policy. Most analysts had already drawn a conclusion that monetary policy had been tried and had been rendered ineffective, mostly due to the presence of other influences. Laurent (1999) provided evidence that such views were widely distorted and that it was a serious mistake not to employ M2. He established that real M2 had a superior performance in predicting real GNP than a number of other competing monetary indicators. The ebbs in the use of monetary policy, however, reversed itself substantially during and in the post global financial and economic crisis. In fact, to stimulate the US economy through one of its worst financial economic crisis in decades, the US Federal government has been resorting to unconventional monetary policies that have raised a series of discussions and analyses, notably concerning the link between money supply and inflation expectation. One such policy that has been very recently announced\(^3\) is referred to as “quantitative easing” whereby the Fed purchases government bonds and other financial assets to increase the money supply, thereby increasing the excess reserves of the banking system, and raising the prices of the financial assets bought.

Variations across major monetary instruments are not uncommon. Though these measures generally fluctuate together, yet the slight variations among them have raised the debate, as well as the quest for, as to which measure of money is the better policy indicator. M1 and M2 are the focus of this paper. Figure 1 depicts three real money measures: the narrower measure of money (M1), and the two broader measures (M2 and M3)\(^4\). As can be seen from the graphs, some variations do indeed exist among these measures, their dynamics are different, their correlations with one another are neither the same nor constant over time, and, finally, it remains to be seen whether M1 or M2 is a better policy tool, or a better measure for the investors and the brokerage firms that are in quest of better predictive measures.

**Figure 1: Growth in M1, M2, and M3**

![Graph showing growth in M1, M2, and M3](image)

This figure shows the growth in M1, M2, and M3 during 1984-2010. M3 series has been discontinued by the Federal Reserve since 2006.

Recent research on equity markets assigns an important role to money aggregates and seeks, among other goals, methods to enhance the predictive powers of such aggregates in forecasting stock prices (returns). Already,\(^4\) The U.S. Fed has ceased to publish the data on M3 since early in 2006.
some measures of money aggregates are considered as indispensable technical indicators in some brokerage houses [see, for example, Baum (2002)]. Considering that the Fed still sets targets for M2 (and M3 until 2006), and on the basis of some theoretical considerations, most researchers today tend to conclude, though unscientifically, that the broader money measures are likely to offer better predictive signals than the narrow money measures, and even the interest-rate-based measures.

This paper contributes to the literature in a few fronts. First, it is one of the first to examine M2, the broader aggregate measure of money supply. It compares the predictive power of M2 in forecasting stock returns with M1, the relatively narrower but more established measure of money supply. Prior studies in the literature have mainly focused on M1 to examine the relationship between money supply and stock returns. Lacking are studies that investigate both M1 and M2 simultaneously and compare the predictive power of these measures in forecasting stock returns. Second, it examines the recent financial crisis in light of these two monetary measures that have been significantly employed to usher the US economy out of recession. These measures are concurrently, at least politically, the subject of some controversies [see, for instance, Chan (2010)]. Third, the data set that is being employed is the most updated data on stock indices and money supply measures that cover not only the financial crisis period of 2007-2009, but some of the events in the post crisis era. Such events, interestingly, are exclusively monetary and are directly and intrinsically related to M1 or M2. For instance, most of the stimulus packages that the Fed has used so far are monetary in nature and involve pumping money into the economy. The most recent example, as was cited above, is “quantitative easing”. We find it very timely and relevant to examine this unique period in light of our methodology and results when examining the relations between the two different measures of money supply and the stock market. Finally, we also believe this paper is the first to examine the relation between M1, M2 and stock indices during and in the post global financial crisis. We have stayed focused on M1 and M2 mainly because research on the predicting power of money supply in the literature is mainly on M1, but economists have been suggesting that M2 is a better measure but research on the relationship between M2 and stock markets has been lacking. We do not include M3 in the analysis since the Fed tends to

---

5 We are highly indebted to one of the reviewers for suggesting and encouraging us to expand the paper and explore its relevance to worthwhile events.
focus more on M2 and more importantly, the Fed has ceased to publish the
data on M3 since early in 2006.

The remaining parts of the paper are organized as follows. Section II
briefly reviews the prior work on money and stock returns. The methodology
and empirical applications then follow, respectively, in sections III and IV.
The last section includes summary and conclusions.

2 - Prior Research

Limiting our review to the post-1970 years, research by a number of
monetary economists establishes the existence of a significant relationship
between changes in the supply of money and measures of stock returns.
Furthermore, stock prices appear to respond to monetary disturbances with
some lags. This phenomenon, which is referred to as monetary portfolio
hypothesis, suggests that a linkage does exist between changes in the supply
of money and changes in the prices of other assets held in an investor’s
portfolio. For example, if there is an increase in money supply, investors
respond to this wealth effect by exchanging money for a variety of assets such
as bonds, stocks, etc., thereby affecting the prices (returns) of such assets.

An early non-econometric test of the monetary portfolio hypothesis
belongs to Palmer (1970) who, using graphs, concluded that changes in
money supply result in changes in stock prices. Subsequently, econometric
studies by Homa and Jaffe (1971), Hamburger and Kochin (1972), Dhakal,
Kandil and Sharma (1993) confirmed a strong linkage between money supply
and stock prices, and also inferred that monetary changes led stock prices.

The implication of the above findings is that investors, using forecast
of changes in money supply, can predict future stock prices and returns. This
conclusion that contradicts the market efficiency hypothesis (EMH) did not,
however, remain unchallenged. Cooper (1974) and Rozeff (1974) argue that
the lead/lag and cross spectra of stock returns and changes in money supply
are consistent with EMH, and that stock returns actually anticipate changes in
the growth of money supply. In their attempts to resolve the contradiction,
Cooper and Rozeff reversed the direction of causality that was already fairly
established by other researchers. This reversal in causality direction, which
was the unintended consequence of providing support for EMH, was later
further extended. For example, Kraft and Kraft (1977) reported that changes
in the S&P price index did affect the narrowly defined money supply, i.e.,
M1. Rogalski and Vinso (1977) and Jones and Uri (1987) concluded that
causality did not appear to go from money supply but rather from stock prices
to money supply. Hence, investors were not able to develop profitable trading rules with information on changes in the money supply.

Recent research also suggests that M2 is not immune to changes in the equity markets. For example, Duca (2000), considering the flow of money between M2 and bond and equity mutual funds, reports that large capital gains or losses in the funds introduce volatility in M2. Even capital gains taxes, he argues, that induce the investors to hold onto gains in the bond/stock market could cause shifts in the M2 velocity. Duca reports the existence of significant causal relations running from mutual fund loads to M2 velocity. His data, in particular, reveals that long-run movements in taxable stock and bond loads lead into velocity changes, whereas only short-run, not long-run, changes in the M2 velocity lead movements in the funds loads.

In contrast with Duca’s work, one could infer a reverse position from the work of Bennett and Sias (2001). They document that money flows, e.g., the difference between uptick and downtick dollar trading volumes, are highly correlated with contemporaneous returns. Also, since money flows exhibit persistence, they are thus good predictors of variations in future stock returns. Money flows -- whether money is flowing into or out of a particular security -- has been used as a technical indicator since the early 1970s. Similar to M2, the popularity of this measure has increased recently. Simply noted, if investors are willing to put money into stock, then the effect on the stock price is fairly clear. Hence causality goes from money to stocks.

Starting in the early 1990’s, the substitutability between M2 and stock and bond funds became increasingly noticeable and, inevitably, confounded the causality analysis further [see, for example, Carlson and Schwarz (1999)]. The substitutability between these variables was demonstrated through the use of straightforward household data. For example, Kennickel, Starr, McCluer, and Suden (1997) document a shift away from bank deposits, i.e. M2, towards mutual funds in the 1990’s [see also Laderman (1997), for a counter argument]. It is also plausible that financial innovation and technology could have influenced this shift, or could have facilitated other linkages among these variables.

Though no consensus at the empirical level has yet been established, recent tests including cointegration and error correction models have enlightened the above controversies substantially. Given the continuing disparity in views, we would be remiss not to conjecture that, in our opinion, a substantial part of the controversies on this topic has arisen due mainly to lack of a unified basis in data coverage, data frequency, model specification, and
the estimation techniques used. It should also be noted that nearly all of the above studies typically employed M1 as a measure of money supply.

In sum, to date, there is a lack of consensus among researchers in explaining the direction of causality between money supply and stock returns. Further lacking in these analyses is also research in determining the predictive superiority of the alternative measures of money, such as M1 and M2. An effort to resolve these points is among the purposes of this paper.

3 – Methodology

The methodology employed in this study is based on Granger’s causality, Johansen and Juselius (1990) cointegration analysis, and Engle and Granger (1987) error-correction representation. These techniques are complemented by tests and procedures advanced by Sims (1972), Pierce and Haugh (1977), and Akaike’s final prediction error (FPE or AIC). The final prediction error (FPE or AIC) test on the choice of the lag structure is considered preferable and is thus mostly used in this paper. It is shown that FPE, particularly under predictive scenarios, performs better, and is well suited for finite samples [see Geweke, Meese and Dent (1983)].

Within the framework of this paper, the Granger causality tests involve the estimation of the following equations:

\[
R_{j,t} = \alpha + \sum_{i=1}^{N_1} \delta_i R_{j,t-i} + \sum_{t=1}^{N_2} \Phi_i \Delta M_{t-i} + \varepsilon_t
\]

\[
\Delta M_t = \beta + \sum_{i=1}^{N_1} \gamma_i \Delta M_{t-i} + \sum_{t=1}^{N_2} \phi_i R_{j,t-i} + \omega_t
\]

where:

- \( R_{j,t} \) is the return on the \( j \)th stock index in period \( t \), and \( \Delta M_t \) is the change in M1 (or M2) in period \( t \),
- \( \varepsilon_t \) and \( \omega_t \) are identically and independently distributed random error terms, and
- \( N_1, N_2, N_3, \) and \( N_4 \) are the length of the lags, appropriately determined in the estimation process by FPE or AIC.

The equations (1) and (2) are intended to establish the causal relations between the change in money aggregates (\( \Delta M1 \) or \( \Delta M2 \)) and the returns on the stock exchanges. We use the change in M1 and M2 instead of the levels because these series are not stationary in their levels, but they are stationary in
their first differences. These results will be revealed in Table 1 at a later point. Our purpose in revisiting (1) and (2) is to resolve the disparity that exists in the prior literature on the causal relations between the various measures of money and stock returns. Under the null hypothesis of no causal relation running from $\Delta M$, to $R_{j,t}$, the coefficients on the lagged values of $\Delta M$ in the first equation should not be statistically significantly different from zero as a group. Similarly, under the null hypothesis of no causal relation running from $R_{j,t}$ to $\Delta M$, the coefficient on the lagged values of stock returns in the second equation should not be statistically significantly different from zero.

Prerequisite to the application of system (1) and (2) is, of course, an examination of the stationary properties of $R_{j,t}$ and $\Delta M$. We will apply augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests in this regard. Most economic variables are I(1) at their levels, and I(0) in their first differences. We anticipate the same to prevail for the series of M1, M2, and prices on the equity exchanges.

In addition to establishing the direction of causality, the second area of our interest is to probe into the nature of the dynamics, both short- and long-term, among changes in the money aggregates and the stock returns. In brief, we would like to, first, establish the long-run co-movements or equilibrium relations between these variables, i.e., whether any cointegrating relations exist, and, then, examine the short-run dynamics among them, if any. We take each in turn.

To check on the existence of cointegration among variables, this paper uses the VAR-based tests of Johansen and Juselius (1990). The procedure is based on the maximum likelihood estimation of the VAR model. From the estimation, two statistics, the trace and the maximum Eigenvalue, are computed to test for the existence of $r$ cointegrating vectors. In particular, the trace statistics tests the null hypothesis of less than or equal $r$ cointegrating vectors against the alternative of more than $r$ cointegrating vectors. The maximum Eigenvalue statistics tests the null hypothesis of $r$ cointegrating vectors against the alternative of $r + 1$ cointegrating vectors.

In general, evidence for co-integration suggests the presence of a long-run relation between the variables. This means that any deviations from this long-run equilibrium will be corrected. Statistically, the presence of co-integration will also rule out non-causality between the variables. As has become evident, long-run equilibrium does not preclude the existence of
short-run dynamics. For example, Miller (1991) notes that if two variables are co-integrated, there must exist temporal causality in the Granger sense between them in at least one direction. This indicates two important channels that might cause the causal changes in the variables. One channel is the response of one variable due to the changes in the other variable, which is viewed as the short-run interactions between the variables. The other channel identifies the adjustment taken by the variables to correct any deviations from an equilibrium path. Therefore, the temporal causality may come from two sources—the sum of the coefficients of the lagged change variables (standard Granger causality test) and/or the coefficient of the lagged error correction term. This implies that the Granger causality tests may be misleading if this second channel is not taken into consideration.

We thus need to address the above concern that requires accounting, simultaneously, for both the short- and the long-term interactions among the money aggregates and the stock returns. We resort, in our third investigation in this regard, to the Engle and Granger (1987) error-correction representation/model (ECM). This methodology that is widely recognized addresses this phenomenon effectively. Within our context, specification of the ECM for the two equations in system (1) and (2) implies:

\[ R_{j,t} = \alpha + \sum_{i=1}^{N_1} \delta_i R_{j,t-i} + \sum_{t=1}^{N_2} \Phi_t \Delta M_{t-i} + \lambda E_{C_{t-1}} + \varepsilon_t \]  \( (3) \)

\[ \Delta M_t = \beta + \sum_{i=1}^{N_1} \gamma_i \Delta M_{t-i} + \sum_{t=1}^{N_4} \phi_t R_{j,t-t} + \kappa E_{C_{t-1}} + \omega_t \]  \( (4) \)

where the EC term is the residuals from the co-integrating regression, \( R_{j,t} \) and \( \Delta M_t \) are I(0), and the lag values of \( N_1 \) through \( N_4 \) are appropriately determined at the estimation level within the context of the specifications in relations (3) and (4).

Relations (3) or (4) provide two channels of causation: short-run, indicated by the joint significance of the coefficients of the lagged causal variables; and long-run, provided by the adjustment coefficients of the lagged deviations from the long-run equilibrium path, i.e., the coefficients of the EC terms. Statistically significant, and less than 1, the coefficients of the EC terms indicate the degree that the dependent variables adjust toward their long-run equilibrium values.
4 - Data and empirical application

The empirical analysis is based on four U. S. equity indices -- S&P500, Dow Jones Industrial, Nasdaq, and Wilshire 5000 composite; and two measures of money supply aggregates -- M1 and M2. The bivariate relation between each of the equity indices and the money aggregates are estimated using weekly data covering the period of Jan 2, 1984 to November 8, 2010. This period includes not only the global financial crisis of 2007-2009, but also about 12 months post the crisis period.

Table 1: Stationary Tests of Stock Exchange Indices and Money Supply Measures M1 and M2

<table>
<thead>
<tr>
<th>Variables</th>
<th>Level</th>
<th>First Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lags</td>
<td>ADF</td>
</tr>
<tr>
<td>Dow Jones</td>
<td>5</td>
<td>-1.92</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(0.632)</td>
</tr>
<tr>
<td>Nasdaq</td>
<td>6</td>
<td>-1.38</td>
</tr>
<tr>
<td></td>
<td>(0.59)</td>
<td>(0.68)</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>5</td>
<td>-1.94</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>Wilshire 5000</td>
<td>7</td>
<td>-1.73</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.53)</td>
</tr>
<tr>
<td>M1</td>
<td>10</td>
<td>-1.93</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.74)</td>
</tr>
<tr>
<td>M2</td>
<td>10</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(0.97)</td>
</tr>
</tbody>
</table>

This table summarizes the statistics of the stationarity tests for the level as well as for the first difference of four stock exchange indices (Dow Jones, Nasdaq, S&P 500, Wilshire 5000) and two measures of money supply (M1 and M2). The Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests are employed to test the null hypothesis of a unit root in the series. The autoregressive (AR) lags are determined by the Akaike Information Criterion (AIC). The p-values are reported in parentheses. All series are initially transformed to natural logarithms. Data are weekly from January 1, 1984 to November 8, 2010.

Our empirical application is based on the methodology that was set forth in section III. The sequence of our application also follows the steps that are outlined in the methodology section. As a prerequisite to Granger analysis, all the variables included in the equations (1) and (2), or (3) and (4), should be stationary. Table 1 reports the results of the augmented Dickey-Fuller and Phillips-Perron tests for all the stock exchange indices (Dow Jones, Nasdaq, S&P 500, Wilshire 5000) and the two measures of money supply
(M1 and M2). The number of lags that are included in each test is based on the minimum Akaike Information Criterion (AIC). For example, for the Dow Jones series, based on AIC, the optimal lag in the unit root model is five in the level and five in the first difference, respectively. The p-values are reported in the parentheses for each of the ADF and PP test statistics. As can be seen from Table 1, all series are nonstationary in their levels. However, after taking the first differences, they all become stationary at the one percent statistical level or below. Therefore, the first difference of these variables will be used in the Granger and ECM equations for causality and long-, short-run analyses.

As noted in the methodology section, the results of the standard Granger causality tests may be misleading if the cointegration between the variables is not considered. We thus examine pairwise cointegration between the variables. Table 2 presents the results of Johansen and Juselius (1990) cointegration tests. Based on initial computed AICs, the orders of lag structure are set equal to four, though higher orders of up to 10 are also examined. Both the trace and the maximum Eigenvalue statistics firmly reject the null hypothesis of no cointegrating vector at one percent level of confidence or below. However, the trace statistics cannot reject the null hypothesis of less than or equal one cointegrating vector nor can the maximum Eigenvalue statistics reject the null hypothesis of one cointegrating vector. Therefore, the result suggests the presence of a unique cointegrating vector between money supply and stock indices returns. The error correction model, therefore, appears more appropriate for the analysis of causal relation between the variables than the standard Granger presentation. However, for the comparison purpose, we estimate and present the results of both models.

To determine the optimal number of lags used in system (1) and (2) or system (3) and (4), we use minimum FPE criterion in an iterative procedure. As an exposition, Table 3 reports the results of the first step of this iterative

---

6 The testing procedure for the ADF and PP tests is applied to the following model: $$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$ where \(\alpha\) is a constant, \(\beta\) is the coefficient on a time trend and \(p\) is the lag order of the autoregressive process.

7 The order of lag in Johansen and Juselius (1990) cointegration test is typically set equal to 4 in the literature. This confirms our initial estimates. We also tried higher order of lag up to 10, but the outcomes were similar.
This table reports the results of cointegration tests based on Johansen and Juselius (1990). Critical values for the test statistics are from Osterwald-Lenum (1992). The 5% CV and the 1% CV are the critical values at the 5 percent and 1 percent level, respectively. All tests include intercept but no trend in CE; and no intercept/trend in VAR. Log transformation of the series is employed throughout. For the trace test, the null hypothesis is that the number of cointegrating vectors is less than or equal to r, against the alternative of higher than r. For the maximum eigen value test, the null hypothesis is that the number of cointegrating vector is r, against the alternative of r+1 cointegrating vectors. Data are weekly from January 1, 1984, to November 8, 2010. *** Significant at the 1 percent level or below.

procedure for only the stock returns. A similar approach is employed for the change in the money supply variables. Additionally, in a second step of this iterative procedure, separate rounds of iterations for other variables are repeated to reach the final optimal values of the lag structures. The outcome of the entire iterative procedure is summarized in Table 4 (see below for more information).
Table 3: One-Dimensional Autoregressive FPEs for the Stock Index Returns

<table>
<thead>
<tr>
<th>Lag</th>
<th>Dow Jones</th>
<th>Nasdaq</th>
<th>S&amp;P 500</th>
<th>Wilshire 5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.7203</td>
<td>1.2188</td>
<td>0.7214</td>
<td>0.6836</td>
</tr>
<tr>
<td>2</td>
<td>0.7200*</td>
<td>1.2183</td>
<td>0.7204*</td>
<td>0.6829*</td>
</tr>
<tr>
<td>3</td>
<td>0.7216</td>
<td>1.2192</td>
<td>0.7218</td>
<td>0.6836</td>
</tr>
<tr>
<td>4</td>
<td>0.7227</td>
<td>1.2180</td>
<td>0.7224</td>
<td>0.6843</td>
</tr>
<tr>
<td>5</td>
<td>0.7232</td>
<td>1.2164*</td>
<td>0.7235</td>
<td>0.6847</td>
</tr>
<tr>
<td>6</td>
<td>0.7244</td>
<td>1.2173</td>
<td>0.7247</td>
<td>0.6855</td>
</tr>
<tr>
<td>7</td>
<td>0.7255</td>
<td>1.2192</td>
<td>0.7258</td>
<td>0.6863</td>
</tr>
<tr>
<td>8</td>
<td>0.7263</td>
<td>1.2211</td>
<td>0.7269</td>
<td>0.6874</td>
</tr>
<tr>
<td>9</td>
<td>0.7267</td>
<td>1.2230</td>
<td>0.7278</td>
<td>0.6884</td>
</tr>
<tr>
<td>10</td>
<td>0.7278</td>
<td>1.2246</td>
<td>0.7285</td>
<td>0.6873</td>
</tr>
</tbody>
</table>

This table reports the Final Predicting Errors (FPEs) of the one-dimensional autoregressive process of four stock index returns. The maximum order of lags was set to ten. Higher lag orders up to twenty were also tried. The outcome was not different. The minimum FPEs for S&P 500 (ΔSP), Nasdaq (ΔNQ), Dow Jones (ΔDJ), and Wilshire 5000 (ΔWI) series in the one-dimensional autoregressive process, indicate optimum lag order of two, five, two, and one in these series, respectively. The (*) indicates the minimum FPE, hence the optimal lag order in each series. All series are in the first difference form (Δ), not in the level since the original levels are not stationary. The data are weekly from January 1, 1984, to November 8, 2010.

Table 3 reports the FPEs of the one-dimensional autoregressive process of only the stock index returns. The maximum order of lags was set to ten. Higher lag orders up to twenty were also tried, however, the outcome was not different. The minimum FPEs for ΔDJ, ΔNQ, ΔSP, and ΔWI series in the one-dimensional autoregressive process indicate optimum lag order of two, five, two, and two in these series, respectively. Holding the order of lag structures of the dependent variables in system (1) and (2), or system (3) and (4), to the ones specified in the first iterative procedure, new FPEs for these relations are then calculated by varying the order of lags of the causal variables from one up to reaching their optimal lag orders and beyond. For example, in the test of whether ΔM1 Granger-causes ΔNQ, the optimum lag order for ΔNQ is held at five as is dictated by its FPE in the first step and as is reported in Table 3. Then the order of the lag of ΔM1 is varied from 1 to n where maximum n is set equal to ten. Corresponding FPEs are computed. The optimal lag order of ΔM1, which corresponds to the minimum of the FPEs that are now computed, is then selected. The optimal lag order of the causal variables is presented in the fourth column in Table 4. Table 4 reports results
of causality tests from two measures of money supply (ΔM1 and ΔM2) to the changes in each stock index (ΔDJ, ΔNQ, ΔSP, ΔWI). Only minimum values of the FPEs and thereby the optimal order of lags of causal variables are included in Table 4.

In the case of ΔM1, as can be seen from Table 4 (Panel A), there is causality from ΔM1 to stock index returns in all cases. The error correction terms are significant in all cases. The results imply that there is both long term relation, and short term causality, between ΔM1 and stock index returns. Further, the size of the error correction coefficients indicates very sluggish speed for the effect of ΔM1 to materialize. These coefficients are within the very low range of -0.0050 to -0.0060, indicating that it will take a while before the full effect of ΔM1 finds its way through the stock market. These findings have important policy implication when the Fed is searching for a policy measure to pull the US economy as fast as possible out of recession and high unemployment rate.

In the ΔM2 cases, the error correction terms are not significant in all cases. However, the F-values are all significant. These suggest that there are short term interactions, but not long term relations between ΔM2 and stock index returns. The policy implications of these findings are interesting and attest to why the stock exchanges and financial policy makers should be more interested in M2. The impact of changes in M2 is short term only and in a stock market wherein long term returns are, in general, often attributed to the evolution of other variables, short –term outcomes are of immediate interest. The same situation applies when the entire financial market is considered, particularly if it is under stress and suffers from a high unemployment rate. Thus, the immediate or short-run impact of any monetary stimulus packages that are employed to invigorate the economy need to be sought through tracing M2, rather than M1. Some of the recent Fed’s monetary policies, for instance its “quantitative easing” to pump in a huge amount of money into the economy could be argued to stem from and are obviously supported by these findings.

To validate system (3) and (4), we have examined a few diagnostic measures. To check the estimated residuals, we have computed the Durbin-Watson (DW) statistics to determine the existence of first order autocorrelation. We also have calculated the Ljung-Box (Q24) statistic to test the existence of serial correlation up to 24 lags in the residual series. We find

---

8 About 9.5 percent in February 2011
Table 4: Analysis of Causality from Money Supply $\Delta M1$ and $\Delta M2$ to Stock Index Returns

<table>
<thead>
<tr>
<th>Dep. variable</th>
<th>Optimal lag of dep. variable</th>
<th>Causal variable</th>
<th>Optimal lag of causal variable</th>
<th>Min. FPE</th>
<th>DW</th>
<th>Q_{24}</th>
<th>ECM term</th>
<th>F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Stock indices and M1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta DJ$</td>
<td>2</td>
<td>$\Delta M1$, ECM1DJ(1)</td>
<td>6</td>
<td>0.717</td>
<td>2.002</td>
<td>25.14</td>
<td>-0.005**</td>
<td>3.44***</td>
</tr>
<tr>
<td>$\Delta NQ$</td>
<td>5</td>
<td>$\Delta M1$, ECM1NQ(1)</td>
<td>2</td>
<td>1.214</td>
<td>2.003</td>
<td>18.07</td>
<td>-0.006**</td>
<td>3.25***</td>
</tr>
<tr>
<td>$\Delta SP$</td>
<td>2</td>
<td>$\Delta M1$, ECM1SP(1)</td>
<td>6</td>
<td>0.718</td>
<td>2.001</td>
<td>21.00</td>
<td>-0.005**</td>
<td>3.90***</td>
</tr>
<tr>
<td>$\Delta WI$</td>
<td>2</td>
<td>$\Delta M1$, ECM1WI(1)</td>
<td>6</td>
<td>0.681</td>
<td>1.996</td>
<td>23.10</td>
<td>-0.005**</td>
<td>4.25***</td>
</tr>
<tr>
<td>Panel B: Stock indices and M2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta DJ$</td>
<td>2</td>
<td>$\Delta M2$, ECM2DJ(1)</td>
<td>2</td>
<td>0.603</td>
<td>2.000</td>
<td>27.31</td>
<td>4.79***</td>
<td></td>
</tr>
<tr>
<td>$\Delta DJ$</td>
<td>2</td>
<td>$\Delta M2$, ECM2DJ(1)</td>
<td>2</td>
<td>0.603</td>
<td>2.000</td>
<td>26.71</td>
<td>-0.002</td>
<td>4.05***</td>
</tr>
<tr>
<td>$\Delta NQ$</td>
<td>5</td>
<td>$\Delta M2$, ECM2NQ(1)</td>
<td>2</td>
<td>1.213</td>
<td>2.004</td>
<td>18.72</td>
<td>3.48***</td>
<td></td>
</tr>
<tr>
<td>$\Delta NQ$</td>
<td>5</td>
<td>$\Delta M2$, ECM2NQ(1)</td>
<td>2</td>
<td>1.213</td>
<td>2.004</td>
<td>18.18</td>
<td>-0.003</td>
<td>3.34***</td>
</tr>
<tr>
<td>$\Delta SP$</td>
<td>2</td>
<td>$\Delta M2$, ECM2SP(1)</td>
<td>2</td>
<td>0.570</td>
<td>1.998</td>
<td>24.02</td>
<td>6.68***</td>
<td></td>
</tr>
<tr>
<td>$\Delta SP$</td>
<td>2</td>
<td>$\Delta M2$, ECM2SP(1)</td>
<td>2</td>
<td>0.571</td>
<td>1.998</td>
<td>23.80</td>
<td>-0.002</td>
<td>5.55***</td>
</tr>
<tr>
<td>$\Delta WI$</td>
<td>2</td>
<td>$\Delta M2$, ECM2NQ(1)</td>
<td>2</td>
<td>0.519</td>
<td>1.993</td>
<td>24.41</td>
<td>0.49***</td>
<td></td>
</tr>
<tr>
<td>$\Delta WI$</td>
<td>2</td>
<td>$\Delta M2$, ECM2NQ(1)</td>
<td>2</td>
<td>0.519</td>
<td>1.993</td>
<td>24.41</td>
<td>-0.002</td>
<td>7.44***</td>
</tr>
</tbody>
</table>

This table reports the results of causality tests for two measures of money supply (causal variables) to stock index returns (dependent variables). The optimum lag orders of the dependent variables are taken from Table 3. Given the lag orders of the dependent variables, the optimum lag orders of the causal variables are determined using the same procedure. Only minimum values of FPEs and thereby the optimal lag orders of the causal variables are included in the table. DW is the Durbin-Watson statistics to examine the first order autocorrelation in the residual series of the models. Q_{24} is the Ljung-Box statistics to examine the serial correlation in the residual series for up to 24 lags. Panel A reports the results for stock indices and M1. Since there is a long-term relationship between M1 and stock indices as shown in Table 2, Error Correction Model (ECM) is used for the analysis. Panel A reports the results for stock indices and M2. Since there is mixed evidence whether M2 is cointegrated with stock indices, both standard Granger causality and ECM are used for the analysis. The t-values of the coefficients of the lagged error correction term are also included. The last column reports the F-values of the null hypothesis of no causality across variable pairs. The data are weekly from January 1, 1984 to November 8, 2010. Significant levels at *, **, and *** indicate 10 percent, 5 percent, and 1 percent, respectively.
that all DW statistics are in the neighborhood of 2, indicating that there is no first order autocorrelation in the residual series in all models. Also, all Q statistics are not able to reject the null hypothesis of no autocorrelation at the 5% level of significance for the 24 lags. Therefore, the results support the validity of our estimated models for $\Delta M_1$, $\Delta M_2$ and the stock index returns.

We now turn to the predictability power of systems (1) through (4). Based on the results from Table 4, we observe that the error correction terms are only significant for $\Delta M_1$ and not for $\Delta M_2$. Therefore, based on our theoretical construct and in light of our empirical results, ECM is appropriate for $\Delta M_1$ and the standard Granger causality is the right theoretical and empirical model for $\Delta M_2$. Given that our interest in the predictability of system (1) through (4) is a practical one and thus lies in the uni-directional causality from $\Delta M_1$ and $\Delta M_2$ to the stock index returns, system (1) for $\Delta M_2$ and system (3) for $\Delta M_1$ are relevant for any further analysis. Also, since our modeling effort thus far appears robust and well behaved, the specifications (1) and (3) are ready to put into further tests to predict the returns on the equity stock exchanges. Two different scenarios are employed: in-sample forecasting (Table 5) and out-of-sample forecasting (Table 6).

An examination of a possible caveat in our evaluation of the forecasting power of $\Delta M_1$ and $\Delta M_2$ is in order. On a purely econometric basis, it could be argued that relation (3) includes one more valuable than relation (1); hence, the comparison may be flawed. Our immediate response is that our approach is based on theoretically sound and empirically supported methodologies. Hence, we do not regard this potential criticism a valid caveat. Having established this fact up front, out of curiosity we estimated the wrong specification for relation (1) by including the ECM term in it. We then followed our estimation and forecasting procedures (see below). The results did not change and as a matter of fact further supported our conclusions9.

Table 5 includes the results of the in-sample forecasting of the stock index returns using $\Delta M_1$ and $\Delta M_2$. In this scenario, the estimation is over the entire sample. To measure the predictability outcome of the estimated models, four criteria are employed: mean square errors (MSE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) The results are mixed and do not show a discernable difference between the predictabilities based on $\Delta M_1$ or $\Delta M_2$.

---

9 These results, tabulated as Tables 5A and 6A, are available from the authors.
As is well known, the crucial test of a model is in its out-of-sample performance, particularly in a dynamic forecasting environment wherein the K-1 period forecasts are kept untouched while forecasting the K-period.

Table 5: Forecasting Stock Index Returns Using $\Delta M1$ and $\Delta M2$ (in-sample forecasting)

<table>
<thead>
<tr>
<th>Change in Indices</th>
<th>MSE $\Delta M1$</th>
<th>MSE $\Delta M2$</th>
<th>RMSE $\Delta M1$</th>
<th>RMSE $\Delta M2$</th>
<th>MAE $\Delta M1$</th>
<th>MAE $\Delta M2$</th>
<th>MAPE $\Delta M1$</th>
<th>MAPE $\Delta M2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dow Jones</td>
<td>0.7103</td>
<td>0.7149</td>
<td>0.0266</td>
<td>0.0267</td>
<td>0.0181</td>
<td>0.0183</td>
<td>130.1741</td>
<td>138.3632</td>
</tr>
<tr>
<td>Nasdaq</td>
<td>1.2000</td>
<td>1.2000</td>
<td>0.0346</td>
<td>0.0346</td>
<td>0.0242</td>
<td>0.0242</td>
<td>143.5701</td>
<td>132.6981</td>
</tr>
<tr>
<td>SP 500</td>
<td>0.7114</td>
<td>0.7133</td>
<td>0.0266</td>
<td>0.0267</td>
<td>0.0182</td>
<td>0.0183</td>
<td>157.3251</td>
<td>156.0836</td>
</tr>
<tr>
<td>Wilshire 5000</td>
<td>0.6739</td>
<td>0.6751</td>
<td>0.0259</td>
<td>0.0259</td>
<td>0.0181</td>
<td>0.0181</td>
<td>184.8739</td>
<td>142.6272</td>
</tr>
</tbody>
</table>

This table compares the results of forecasting stock index returns using $\Delta M1$ and $\Delta M2$. Based on the results in table 4, Error Correction Model is used to forecast with $\Delta M1$, while standard Granger causality is used to forecast with $\Delta M2$. The models are estimated using the entire sample. Mean square errors (MSE), Root mean square errors (RMSE), Mean absolute errors (MAE), and Mean absolute percentage errors (MAPE) are the criteria for comparison. The data are weekly from January 1, 1984 to November 8, 2010.

outcomes. To test the robustness of the estimated models and determine the influence of $\Delta M1$ vis-à-vis $\Delta M2$ on the stock market returns, we conduct a series of out-of-sample dynamic forecasts. In brief, the models are estimated using N-K observations, where N is the total number of observations; K is the length of the forecasting windows. Three different windows are explored: four weeks, eight weeks, and twelve weeks. After re-estimating the coefficients in each N-K cases, we dynamically forecast the stock index returns in the next K periods. For instance, for the four-week forecast ahead, we first use N-4 observations to estimate the coefficients of the models and then forecast the stock index returns for the next four-week period after the N-4 observations. The difference between the predicted and the actual values is the error. MSE, RMSE, MAE, and MAPE are then computed and used as criteria for comparison.

The results are reported in table 6. There is a conclusive evidence that $\Delta M2$ is a better candidate for forecasting stock index returns. In almost
all cases, MSE, RMSE, MAE, and MAPE for $\Delta M2$ are lower than those for $\Delta M1$. The policy implications of these results, as were briefly pointed out

Table 6: Dynamic Out-of-Sample Forecasting of Stock Index Returns Using $\Delta M1$ and $\Delta M2$

<table>
<thead>
<tr>
<th>Change in Indices</th>
<th>MSE 4-week forecast ahead</th>
<th>MSE 8-week forecast ahead</th>
<th>MSE 12-week forecast ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta M1$</td>
<td>$\Delta M2$</td>
<td>$\Delta M1$</td>
</tr>
<tr>
<td>Dow Jones</td>
<td>0.1903</td>
<td>0.1193</td>
<td>0.2564</td>
</tr>
<tr>
<td>Nasdaq</td>
<td>0.3888</td>
<td>0.3230</td>
<td>0.5025</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>0.3287</td>
<td>0.2796</td>
<td>0.3423</td>
</tr>
<tr>
<td>Wilshire5000</td>
<td>0.3551</td>
<td>0.3023</td>
<td>0.3978</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change in Indices</th>
<th>RMSE 4-week forecast ahead</th>
<th>RMSE 8-week forecast ahead</th>
<th>RMSE 12-week forecast ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta M1$</td>
<td>$\Delta M2$</td>
<td>$\Delta M1$</td>
</tr>
<tr>
<td>Dow Jones</td>
<td>0.0138</td>
<td>0.0109</td>
<td>0.0160</td>
</tr>
<tr>
<td>Nasdaq</td>
<td>0.0197</td>
<td>0.0180</td>
<td>0.0224</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>0.0181</td>
<td>0.0167</td>
<td>0.0185</td>
</tr>
<tr>
<td>Wilshire5000</td>
<td>0.0188</td>
<td>0.0174</td>
<td>0.0199</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change in Indices</th>
<th>MAPE 4-week forecast ahead</th>
<th>MAPE 8-week forecast ahead</th>
<th>MAPE 12-week forecast ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta M1$</td>
<td>$\Delta M2$</td>
<td>$\Delta M1$</td>
</tr>
<tr>
<td>Dow Jones</td>
<td>1.3493</td>
<td>1.4351</td>
<td>1.3215</td>
</tr>
<tr>
<td>Nasdaq</td>
<td>0.6438</td>
<td>0.5625</td>
<td>0.9736</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>2.3872</td>
<td>2.5567</td>
<td>1.9451</td>
</tr>
<tr>
<td>Wilshire5000</td>
<td>2.2562</td>
<td>2.6679</td>
<td>2.0789</td>
</tr>
</tbody>
</table>

This table compares the results of dynamic out-of-sample forecasting stock index returns using $\Delta M1$ and $\Delta M2$. Error Correction Model is used for $\Delta M1$ and standard Granger causality presentation is used for $\Delta M2$. The original estimated parameters of the equations are reported in Table 4. The forecasting process is purely dynamic, i.e., no adjustments in any one of the prior forecasts are allowed. The model is re-estimated using N-K observations, where N is the total number of observations, K is the length of the forecasting windows. Three different windows are explored: four-week, eight-week, and twelve-week. After estimating the new sets of coefficients, we dynamically forecast stock index returns in the next periods. For instance, for the four-week forecast ahead, we first use N-4 observations to estimate the coefficients and

10 As was pointed out earlier, out of curiosity, we also use the ECM term in the $\Delta M2$ relation, i.e., the wrong model, instead of the standard Granger causality model (the correct model). The results are unchanged and in all cases, MSE, RMSE, MAE, and MAPE for $\Delta M2$ are lower than those for $\Delta M1$. 129
then forecast stock index returns in the next period after the N-4 observations. The difference between the predicted and the actual values is the error. Mean squared errors (MSE), Root mean square errors (RMSE), Mean absolute errors (MAE), and Mean absolute percentage errors (MAPE) are then computed and used as criteria for comparison. The data are weekly from January 1, 1984 to November 8, 2010.

earlier, suggest that M2 is a more effective measure than M1, both in short-term prediction of the stock market and as a monetary policy tool when immediate or near term results are sought. In this regard, though we are not aware of the internal analyses and deliberations of the U.S Federal government in the selection and design of the stimulus packages that so far have been released to bring about some degree of recovery at the highest speed, it appears that the Fed has been well aware of the relative significance of M2 and thus has acted properly in its choice of monetary instruments. Nearly all of the Fed’s policies have been relying on some direct or indirect monetary measures to turn around the economy, particularly targeting growth and unemployment. These are two very pressing variables, both economically and politically, in any economy that finds itself in a recession. The U.S. Fed “quantitative easing” that was mentioned earlier could be cited as an example again at this point. While long term effects are acknowledged and are viewed desirable, reaching short term results often appear to override the long term ones. Hence, M2, rather than M1, seems to have rendered itself as a more preferred policy measure.

6 - Summary and conclusion

In this paper we have demonstrated that causal relations between stock returns and money supply do exist. Using cointegration and error correction representation of four U.S. equity exchange series (S&P500, Dow Jones Industrial, Nasdaq Industrial, and Wilshire 5000 composite) and two aggregate measures of money supply (M1 and M2), we have established that these relations enjoy long-term equilibrium but are different in specifications and forecasting power. We have considered short-run dynamics as well, and have found that they could be characterized as departures from the long-run equilibrium.

Within the above framework, we have estimated a theoretically correct and empirically robust and well-behaved model for ΔM1 and ΔM2 that conjecture, surprisingly, different specifications. We then have put this model to test to check the predictive power of the two money aggregates
(ΔM1 and ΔM2). Based on a set of in- and out-of-sample forecast experiments, four measures (MSE, RMSE, MAE, and MAPE) for both ΔM1 and ΔM2 are calculated. The results overwhelmingly indicate that ΔM2 is a better predictive measure and hence a superior indicator than ΔM1.

Our results also support a number of other related findings that together may have some bearing on the choice and use of money aggregates as measures of policy choice, particularly in an economy that is in recession and is seeking immediate or near term turnaround. In this regard, we have demonstrated that the US Fed’s policies of using stimulus packages or its very recent policy of resorting to “quantitative easing” are in line and are well supported by our findings. Our results provide conclusive evidence that ΔM2 is a more effective measure than ΔM1, both as a short-term predictor and as a monetary policy tool when near term results are sought.

Our results also support a few strand of breakthroughs in the literature on the global markets. For example, Bahmani-Oskooee and Chomsisengphet (2002) demonstrate that M2 rather than M1 has a long-run relationship with income and interest rate in many countries. Thus, it could be argued that ΔM2 possesses, or draws in, the parallel effects of income and interest rate in predicting equity returns. Though no formal test on this point is performed, this position adds to the depth of ΔM2 in terms of its economic and financial significance. There are other rationales for considering ΔM2 (see Mehra, 1998). The potentials of ΔM2 as a predictive indicator in other economic and financial domains are not yet fully examined. In this regard, future research is well warranted.

References


Johansen, Soren and Katarina Juselius, 1990. Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the


