“Forecasting US GDP: Assessing the Quantity Theory of Money as a Device”

By Daphne Miyagawa Chang

Honors Thesis

SUNY New Paltz

December 15, 2021

“Forecasting US GDP: Assessing the Quantity Theory of Money as a Device”

Abstract:

We revisit Friedman’s Quantity of Money Theory and analyze the ability to forecast US real GDP using lagged M2 money supply. The Granger Causality Test indicates that M2 money supply Granger causes real GDP. However, we find that the causality is one-way, from M2 to real GDP, as real GDP fails to Granger cause M2. We surveyed a range of ARIMAX models to ascertain whether M2 money supply still forecasts economic activity, both before and during the global COVID pandemic. While economic theory suggests that money is neutral, we found that hybrid models that combine both standard ARIMA parameters plus M2 money supply growth as an exogenous regressor(s) actually improve forecast accuracy. Both before, during, and in the full sample, a hybrid model combining money supply and autoregressive parameters outperform atheoretical Box Jenkins models as well as the naïve model. Our results suggest that, while money supply has fallen off the radar screen of many economists, employing M2 in a hybrid model still improves forecasting ability.

Key words: accounting, Quantity of Money Theory, forecasting, economics, M2, money supply, monetary theory, Granger causality, ARIMA, Box-Jenkins

JEL Classification: B22, C22, E17, E51

1. Introduction:

Herein, we examine the forecasting prowess of models that take money supply growth into consideration vs. standard atheoretical ARIMA models and a naïve model of just the sample mean. Based on Friedman’s version of the Quantity Theory of Money (QTM), we are led to believe that money supply is the chief determinant of (real) Gross Domestic Product (GDP). However, economic literature has long held that money is neutral in the long run. To better understand the potential for money supply as a forecasting device, we constructed a forecast competition study to assess the viability of models that include a measure of money growth vs. models that are purely atheoretical.

Friedman had strong beliefs that M2 money supply is stable and a good indicator to predict future output growth. M2 money supply includes M1 (currency, cash, coins held by the public) and saving deposits and small-time deposits, as well as shares in retail money market mutual funds.

According to the quantity of money formula, the stability of velocity will affect our ability to forecast real GDP with money supply alone. This is shown by the equation below:

$$Money * Velocity = Price * Quantity = GDP$$

This equation relies on the key assumption that velocity is constant. Empirically, we observe that this is far from the case for M1 money supply. In fact, M2 money supply is more stable. We will rely on the "well-behaved" variation in M2 and use that aggregate for forecasting.

The main hypothesis we are testing is whether if M2 money supply helps us to forecast real GDP. In other words, if the lags of M2 money supply help in predicting real GDP. We employ the Granger Causality Test in order to conduct this analysis.[1]

Our test results indicate that M2 money supply Granger causes real GDP. However, the causality is one-way, as real GDP fails to Granger cause M2 money supply.

Given these results, we conduct a forecast competition study using EViews to see if M2 money supply can forecast real GDP better than a naïve model, vs. standard atheoretical ARIMA models, and in hybrid autoregressive distributed lag (ARDL) models that employ both autoregressive terms with lags of M2 growth as exogenous regressors.

The results from our forecasting models do, in fact, suggest that hybrid models that employ M2 as an exogenous regressor outperform the naïve and atheoretical ARIMA models.

2. Literature Review:

In this section, we shall provide a survey of the relevant literature. We focus on the contributions that come from the Monetarist framework of Milton Friedman and his predecessors and highlight some additional more-recent contributions in this area.

Friedman objected to discretion in monetary policy. In Friedman (1992), he argued that discretion "meant continual and unpredictable shifts" in policy. Specifically, he believed that the role of monetary authorities is to stabilize the economy, and that discretion does the complete contrary. He also wished that monetary authorities be objective and not rely on personal perspectives.[2]

In addition, Friedman also favored a policy rule where the M2 money supply would grow between 3 to 5 percent annually, signifying stable growth (1992).[3] In an interview with Professor James Forest of SUNY New Paltz, conducted in 1998 when he was a Market Analyst at Standard & Poor's, Friedman stated that M2 money supply is the right indicator to measure long-term growth. In addition, he said, "I believe the only function the Federal Reserve has is to keep M2 growing slowly."[4]

Friedman also held controversial views on what he believed were necessary changes to the U.S. monetary policy. According to Lothian (2009), Friedman also called for the imposition of 100

[1] See: Granger's *Testing for Causality: A Personal Viewpoint* (1980), page 48
Also Brook's *Introductory Econometric for Finance, 3rd Edition* (2014), page 335-350
[2] See: Friedman *A Program For Monetary Stability* (1992), page 85
[3] ibid, page 91
[4] See: Dr. James Forest's interview with Milton Friedman (1998)
(https://www.youtube.com/watch?v=MiID2cU1qI8)

percent required reserves on depository institutions and wanted the Fed to maintain money supply in a constant growth state.[5]

Ip and Whitehouse (2006) suggest that Friedman and Phelps were proven correct when Paul Volcker became the Federal Reserve Chairman in 1979 and increased interest rates to double-digits, bringing the economy into a deep recession.[6] This ultimately reduced inflation from 6.13% in 1982 to 1.9% in 1986.[7] The unemployment rate fell from 9.7% in 1982 to 7% in 1986, which gradually stabilized to 5.3% in 1989, just slightly above the natural unemployment rate.[8]

According to Walsh (2004), Paul Volcker's decisive action in 1979 to adopt a money supply targeting framework at the Fed stabilized the economy and ultimately helped the US maintain low inflation rates. These lower rates of inflation have persisted ever since the Fed began its modern era of prioritizing inflation in its official dual mandate.

Research in monetary economics has been mixed, some suggest money does affect future output. According to Sims (1972), "money can be treated as exogeneous in a regression of GNP on current and past money.[9] On the other hand, additional results suggest that money is neutral in the long run. In Eichenbaum and Singleton (1986), the results of their research suggest that "lagged values of the monetary growth rate are not helpful in predicting the current and future growth rates of output, after conditioning on the other variables in the VARs."[10]

Becketti and Morris (1992) suggest that it is difficult to draw a concrete conclusion, as the results depend on "variables used to measure money and output, on the frequency of observation of the data (monthly or quarterly), on the specification of the Granger test regression, and on the sample used in the regression."[11] Ultimately, M2 money supply is helpful in forecasting future economic activity when the period of change in the Federal Reserve in the 1980s is excluded from studies. Ironically, this is the period when Volcker was targeting money supply. In contrast to Becketti and Morris (1992), we included the period of change in the Federal Reserve in in the 1980s in our forecast models and found that M2 money supply is still useful in forecasting future economic activity.

In Brand, Reimers, and Seitz (2003), it is shown that narrow money (M1) has important indicator properties with respect to real GDP growth. M1 outperforms the yield spread in predicting GDP.[12] Despite this, these findings on European data contradict that of the U.S. In Hamilton and Kim (2002), their research concluded that M1 is not useful in predicting GDP growth.[13] We can conclude that narrow money is a useful predictor of economic output in the euro area but we feel the jury is still out with respect to the US.

---

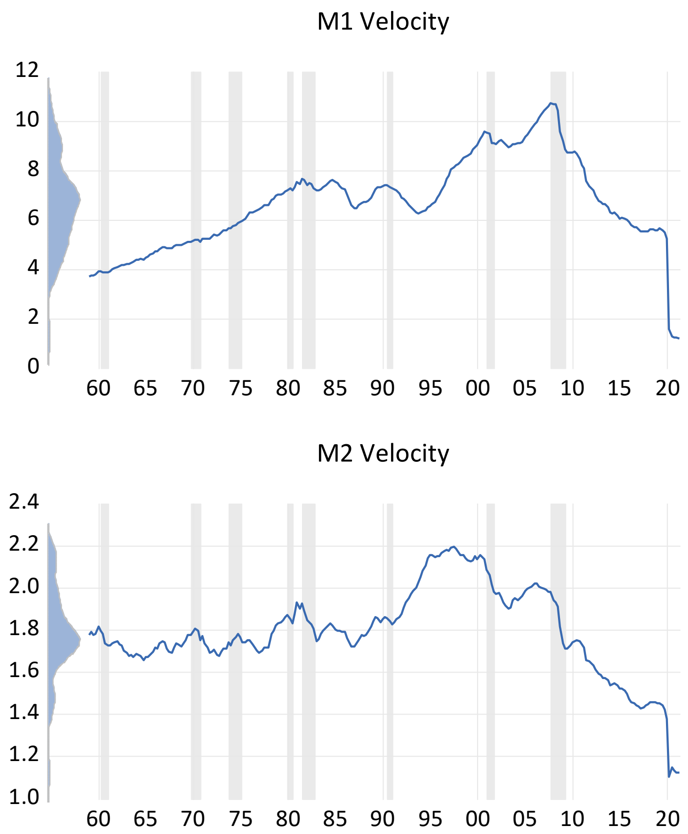
[5] See: Lothian's journal *Milton Friedman's Monetary Economics and the Quantity-Theory Tradition* (2009), page 1092
[6] See: Ip and Whitehouse's *How Milton Friedman Changed Economics, Policy and Markets* (2006), page 2-3
[7] Extracted from FRED's *Inflation, consumer prices for the United States (FPCPITOTLZGUSA)*
[8] Extracted from FRED's *Unemployment Rate (UNRATE)*
[9] See: Sims' *Money, Income, and Causality* (1972), page 550
[10] See: Eichenbaum and Singleton's *Do Equilibrium Real Business Cycle Theories Explain Post-war U.S. Business Cycles?* (1986), page 131-132
[11] See: Becketti and Morris' *Does Money Still Forecast Economic Activity?* (1992), page 68
[12] See: Brand, Reimers, Seitz's *Forecasting Real GDP: What Role for Narrow Money?* (2003), page 26
[13] See: Hamilton and Kim's *A Reexamination of the Predictability of Economic Activity Using the Yield Spread*, page 340

In Geweke (1986), annual U.S. data and postwar monthly data support "structural superneutrality of money with respect to output and the real rate of return".[14] This is counter to what was suggested in the aforementioned study by Becketti and Morris (1992) who suggest that frequency of observation would significantly affect results. Geweke also finds that his interpretation of the "empirical evidence on superneutrality is not affected by variable construction" (1986).[15]

3. Data and Methodology:

Given Friedman's view that M2 is the proper monetary aggregate due to the behavior of its velocity, we decided to use M2 money supply in our models due to the greater stability relative to M1 money supply. Figure 2 confirms that M1 is unstable relative to M2. This is important because the quantity of money identity equation, MV = PQ, relies on a stable velocity of the monetary aggregate. If money velocity is well-behaved, money supply growth can be used as a stand-alone predictor of output. If M2 was found to be seriously lacking in stability, some measure of M2 * estimated velocity would need to be used. This is because velocity is unobservable in the real world. Only after computing real GDP would we be able to use the QTM identity to back out velocity.

Figure 1. M1 Velocity vs M2 Velocity



[14] See: *The Superneutrality of Money in the United States: An Interpretation of the Evidence*, page 1
[15] Ibid, page 19

Table 1. – Descriptive Statistics on Money Velocity

	M1 Velocity	M2 Velocity
Mean	6.65	1.79
Median	6.63	1.77
Maximum	10.70	2.19
Minimum	1.19	1.10
Std. Dev.	1.85	0.21
Skewness	-0.14	-0.40
Kurtosis	3.19	3.97
Jarque-Bera	1.16	16.32
Probability	0.56	0.00
Sum	1661.84	447.31
Sum Sq. Dev.	853.68	10.70
Observations	250	250

According to table 2, we can make comparisons between descriptive statistics on money velocity between M1 and M2 money supply. Standard deviation denotes the relative stability of money velocity. M1 velocity has a standard deviation is 1.85, while M2 velocity is just 0.21. This confirms that Friedman's belief that M2 velocity is well-behaved, at least in a relative sense. Our estimation sample runs from the year 1959 to the year 2016.

Visually, we can observe the relative stability of M2 velocity compared to that of M1 in figure 2. M2 velocity is consistently lower than that of M1 velocity, ranging between 1.10 and 2.19, while M1 velocity ranged from 1.19 to 10.70. We should, however, note the apparent structural break in M2 velocity during the COVID-19 period. In both charts, we observe a sharp drop in money velocity.

4. Preliminary Analysis: Granger Causality Test

We used a vector autoregressive (VAR) test to identify the appropriate lag length for the Granger Causality test from the log transformed RGDP & M2 series. The model suggests that the appropriate lag length, according to Schwartz's Information Criteria, is three periods. Then, we use the Granger Causality test through EViews to examine whether if money growth helps forecast GDP growth. In addition, we utilize the Johansen Cointegration Test to aid with the specification of the Granger Causality Test. The Johansen Cointegration Test indicates no cointegration at the 0.05 significance level in both the Trace Test and the Maximum Eigenvalue Test.

Table 2. Granger Causality Test

Pairwise Granger Causality Tests
Sample: 1959Q2 2021Q2
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Prob.
M2_PCH does not Granger Cause GDP_CHG	246	14.5576	1.00E-08
GDP_CHG does not Granger Cause M2_PCH		1.09146	0.3534

The Granger Causality test results are contained in table 2. The tests were conducted using EViews version 12. We find that the probability of falsely rejecting the null hypothesis (H_0) is near zero, 1.00E-08. The F-statistic is also evidence against the null hypothesis; the bigger the F-statistic, the more confident we are in rejecting the null. Thus, we can conclude that M2 money supply Granger-causes real GDP with a high degree of statistical significance. However, real GDP does not Granger-cause M2 money supply. The relationship is a clear one-way causality from M2 money supply to real GDP.

In the following section, we shall put this evidence to use by building an array of forecasting models, both atheoretical and theoretical. From these models we will proceed by estimating over our estimation period and predict results over our forecast horizon.

5. Estimation Results:

In this section we estimate parameters for a set of alternative forecasting models. The set of models includes the naïve model plus purely atheoretical time series models based on standard Box-Jenkins (ARIMA) methods. The included ARIMA models used were the first-order AR(1) and MA(1) models. We also augment the set of models with a purely theoretical model with a single exogenous predictor, M2. Finally, we formulate several hybrid models of the autoregressive distributive lag with exogenous regressor class (ARDL-X).

Additionally, the naïve model (regression on a constant mean), M2, and several ARDL models were used to forecast economic activity in 2019 (pre-COVID), 2020 (COVID Peak), and 2021 (present time). The dependent variable for all models is real GDP. and independent variable is lags of dependent variable (AR) in the ARIMA models. In the ARDL models, lagged M2 is also a predictor.

The set of model equations are given below. Note that: y_t = quarterly real GDP growth; μ = a parameter estimate of the conditional mean or intercept; ε_t = the error term; φ = is the first-order autoregressive parameter; X = quarterly percentage change in M2 money supply; θ_i = moving average term; β = coefficient on M2 money supply.

1. Naïve model:

$$y_t = \mu + \varepsilon_t$$

2. AR(1):

$$y_t = \mu + \varphi * y_{t-1} + \varepsilon_t$$

3. MA(1):

$$y_t = \mu + \theta_i * \varepsilon_{t-1} + \varepsilon_t$$

4. M2 factor model:

$$y_t = \mu + \beta * X_{t-1} + \varepsilon_t$$

5. ARDL-X (1,1):

$$y_t = \mu + \varphi * y_{t-1} + \beta * X_{t-1} + \varepsilon_t$$

6. ARDL-X (2,1):

$$y_t = \mu + \varphi_1 * y_{t-1} + \varphi_2 * y_{t-2} \ \beta * X_{t-1} + \varepsilon_t$$

Estimation results are displayed in Table 3. We see that the adjusted R-squared ranges from 0, for the naïve model, to 14.6% for the hybrid 2 model in equation 6. This demonstrates quite a range in explanatory prowess across these models. However, it should be noted that even the model with the greatest explanatory power still leaves just under 85% of the total variation left unexplained. The model with M2 as a stand-alone predictor explains only 3.4% of variation in real GDP, while the combination of M2 and a single AR(1) term explains 10.9% of total variation.

In terms of significance, the estimates for one-period-lagged M2 are highly significant in all models that contain this parameter. The lagged M2 parameter ranges from a low of 0.196, in equation 4, to 0.399 in equation 6. The F-statistic in equations 2 to 6 are all statistically significant, indicating that these models are preferable to the naïve model in equation 1. We note that it may be possible to achieve greater explanatory power by increasing the lag order for both the Box-Jenkins (ARMA) components, as well as on the lags of the exogenous regressor, M2. However, the point of the study is to better understand overall forecasting benefits of employing M2 in a parsimonious out-of-sample forecasting model, as opposed to achieving the best in-sample fit.

Table 3. Estimation Results

Eq Name:	NAIVE	AR1	MA1	M2	HYBRID1	HYBRID2
Eq. #:	1	2	3	4	5	6
C	0.007649	0.005297	0.007654	0.004260	0.002481	-0.001649
	[13.8473]**	[7.4185]**	[11.8016]**	[3.4951]**	[2.0081]*	[-1.0743]
GDP_CHG(-1)		0.300508	0.122062		0.279560	0.122062
		[4.7883]**	[2.0050]*		[4.4998]**	[2.0050]*
M2_PCH(-1)				0.195781	0.174941	0.398860
				[2.9993]**	[2.7832]**	[6.4800]**
GDP_CHG(-2)						0.165870
						[2.7973]**
Observations:	232	231	232	230	230	248
Adj. R-squared:	0.0000	0.087041	0.055655	0.033737	0.108961	0.146448
F-statistic:	NA	22.9281	7.8070	8.9955	15.0017	15.1263
Prob(F-stat):	NA	0.0000	0.0005	0.0030	0.0000	0.0000

Table 3. displays the estimation output for our set of models. C refers to a constant, while GDP_CHG(-1) refers to change in GDP from a quarter ago. Naturally, M2_PCH(-1) signifies percentage change in M2 money supply from a quarter ago, and GDP_CHG(-2) signifies change in GDP from two quarters ago. The MA1 (moving average) model performs the best as it achieved the lowest adjusted R-squared and F-statistic.

6. Forecasting Results

We divided our forecasting program into three categories—the full sample, from Q1 of 2017 to Q2 of 2021 (present); Pre-COVID, which includes a sample size dating from Q1 of 2017 to Q4 of 2019. Lastly, we forecast the COVID period, lasting from Q1 of 2020 to Q2 of 2021.

Tables 4. Forecasting Results

Panel A. Full Sample 2017Q1-2021Q2

Observations	Naïve	AR(1)	MA(1)	M2	Hybrid 1	Hybrid 2
Root Mean Squared Error	0.0285	0.032	0.0308	0.0262	0.0289	**0.0252**
Mean Absolute Error	0.0126	0.0142	0.014	0.011	0.0131	**0.0109**
Mean Abs. Percent Error	57.2669	58.9375	60.9627	**48.1495**	56.1245	48.3478
Theil Inequality Coef.	0.7815	0.8079	0.8005	0.6655	0.7449	**0.6092**
Bias Proportion	0.0083	**0.0024**	0.0045	0.0218	0.0096	0.0249
Variance Proportion	NA	0.3866	0.503	0.7667	0.5945	0.6158
Covariance Proportion	NA	0.611	0.4926	0.2116	0.3959	0.3593

Panel B. Pre-COVID 2017Q1-2019Q4

Observations	Naïve	AR(1)	MA(1)	M2	Hybrid 1	Hybrid 2
Root Mean Squared Error	0.0023	0.002	0.0021	0.0021	**0.002**	0.0029
Mean Absolute Error	0.0018	**0.0015**	0.0016	0.0018	0.0017	0.0024
Mean Abs. Percent Error	44.9963	37.2286	39.8844	38.2034	**35.5477**	39.4531
Theil Inequality Coef.	0.1652	**0.145**	0.1502	0.1621	0.153	0.2455
Bias Proportion	0.339	0.2072	0.2914	0.021	**0.0005**	0.2315
Variance Proportion	NA	0.45	0.5337	0.3592	0.3164	0.0451
Covariance Proportion	NA	0.3428	0.1749	0.6197	0.6831	0.7234

Panel C. COVID Period 2020Q1-2021Q2

Observations	Naïve	AR(1)	MA(1)	M2	Hybrid 1	Hybrid 2
Root Mean Squared Error	0.0493	0.0554	0.0533	0.0452	0.05	**0.0434**
Mean Absolute Error	0.0342	0.0395	0.0388	0.0294	0.0358	**0.0278**
Mean Abs. Percent Error	81.8082	102.3553	103.1191	68.0416	97.278	**66.1371**
Theil Inequality Coef.	0.8686	0.8554	0.8581	0.6997	0.7835	**0.6248**
Bias Proportion	0.0105	**0.0028**	0.0054	0.0588	0.0282	0.114
Variance Proportion	NA	0.3858	0.5011	0.8404	0.6152	0.7706
Covariance Proportion	NA	0.6114	0.4935	0.1009	0.3566	0.1154

We opted to use various model benchmarks to serve as a set of alternative forecasts to compare with results from models that contain the theoretically implied money supply component. The set of benchmarks includes the naïve model, AR(1), MA(1), M2 money supply, Hybrid 1, and Hybrid 2 models. The hybrid models combine autoregressive components with lagged M2 in standard autoregressive distributed lag (ARDL-X) form with money supply as the exogenous predictor. The naïve model simply produces forecasts of real GDP based on the estimation period average change.

According to tables 4., we can see that Hybrid 2 model performs the best in the full sample and the COVID period sample, with one distributed lag term of M2 money supply percentage change, and two autoregressive terms. The hybrid 1 model performs the best in the Pre-COVID sample, with one distributed lag term and one autoregressive term.

Our main contribution is that M2 money supply helps us forecast GDP 1 period (quarter) ahead. This improves forecast accuracy compared to strictly atheoretical naïve and simple ARIMA models. Hybrid models that include money supply also outperform forecasts from the stand-alone M2 model. In short, the inclusion of money supply improves forecasting abilities of a single-equation prediction model for real GDP. As per tables 4., we can observe that our Hybrid 2 model is overall the most accurate, with four forecast measures (root mean squared error, mean absolute error, mean absolute percent error, and Theil inequality coefficient) being the lowest value in our forecast of the COVID period.

7. Conclusion:

In conclusion, M2 money supply can increase the accuracy of ARIMA model forecasts. Specifically, money supply helps us forecast real GDP one quarter ahead, improving forecasting over both the naïve and simple ARIMA models. Hybrid models that include money supply out-perform models that do not include money supply, in out-of-sample forecast comparisons.

In order to determine the amount of improvement in accuracy, we determine it through the percentage change between naïve forecast and the “best” model (Hybrid 2 in this case). Although we experience certain limitations during our research, we discovered that money supply is still relevant in increasing the accuracy of forecast models in macroeconomic forecasting.

However, we were only able to obtain a small set of Box-Jenkins models to perform forecasts. If we could perform more types of forecasts, we would have more confidence in determining the best forecast model and it could be different from question 5 in which the consistent model is Hybrid 2. Moreover, attempts to forecast real GDP using ARIMA(X models would be more accurate if we included lagged money supply as an exogenous regressor to reduce the means squared error further and therefore provide more accurate forecasts.

One way the study could be expanded to strengthen our conclusions is to increase the number of alternative models to further consider whether Hybrid 2 is the most precise forecast model out of all alternatives.

References

Becketti, S. and Morris, C. (1992) ‘Does Money Still Forecast Economic Activity?’ *Economic Review – Federal Reserve Bank of Kansas City; Fourth Quarter 1992,* 68

Brand, C., Reimers H. E., and Seitz F. (2003) ‘Forecasting Real GDP: What Role for Narrow Money?’ *European Central Bank, Working Paper Series No. 254,* 26

Brooks, C. (2014) ‘Introductory Econometrics for Finance, 3rd Edition’ *Cambridge University Press,* 335-350

Eichenbaum M and Singleton K. J. (1986) ‘Do Equilibrium Real Business Cycle Theories Explain Postwar U.S. Business Cycles?’ *NBER Macroeconomics Annual, 1986, Vol. 1,* 131-132

Federal Reserve Bank of St. Louis, NBER based Recession Indicators for the United States from the Period following the Peak through the Trough [USREC], retrieved from FRED, *Federal Reserve Bank of St. Louis*; https://fred.stlouisfed.org/series/USREC, November 23, 2021.

Friedman, M. (1992) ‘A Program for Monetary Stability,’ *Fordham University Press*, 85-91.

Geweke, J. (1986) ‘The Superneutrality of Money in the United States: An Interpretation of the Evidence,’ *Econometrica Vol. 54, No. 1*, 1-19

Granger, C. W. J. (1980) ‘Testing for Causality: A Personal Viewpoint,’ *Journal of Economic Dynamics and Control, 2*. 48

Hamilton, J. D. and Kim, D. H. (2002), ‘A Reexamination of the Predictability of Economic Activity Using the Yield Spread,’ *Journal of Money, Credit and Banking, May 20092, Vol. 34, No. 2,* 340

Ip, G. and Whitehouse M. (2006) ‘How Milton Friedman Changed Economics, Policy and Markets,’ *The Wall Street Journal*, 2-3.

Lothian, J. R. (2009) ‘Milton Friedman’s Monetary Economics and the Quantity-Theory Tradition,’ *Journal of International Money and Finance 28,* 1092.

U.S. Bureau of Labor Statistics, Unemployment Rate [UNRATE], retrieved from FRED, *Federal Reserve Bank of St. Louis*; https://fred.stlouisfed.org/series/UNRATE, November 23, 2021.

Sims, C. A. (1972) ‘Money, Income, and Causality,’ *The American Economic Review, Sep. 1972, Vol. 62, No. 4*, 550

World Bank, Inflation, consumer prices for the United States [FPCPITOTLZGUSA], retrieved from FRED, *Federal Reserve Bank of St. Louis*; https://fred.stlouisfed.org/series/FPCPITOTLZGUSA, November 24, 2021.