Compartmentalising Gold Prices

Rohnn Sanderson

Abstract

Deriving a functional form for a series of prices over time is difficult. It is common to assume some linearly estimable form for prediction purposes. While this can produce accurate short run forecasts it fails to identify longer trends and patterns that may exist in financial data. Particularly troublesome is the potential for chaotic behaviour which can look like standard autocorrelation. Also, components of a price's behaviour may not be linear or may be unable to be structured well in a stationary series. Recently, more research has been devoted to whether or not a series of prices exhibits deterministic behaviour, instead of some type of Brownian Motion (regular or fractal). This research suggests that some time series data may pass typical tests for randomness where randomness does not exist. Given the breadth of current research, the most logical and reasonable beginning assumption for modeling a time series is that data probably exhibit both deterministic and random components. This paper will make use of the techniques of spectral analysis and the Hurst Exponent to measure the level of long-run dependence in the price data of gold. This technique will allow for the separation and quantification of how large the deterministic and random components of gold prices are.

Keywords: Dynamic Systems, Hurst Exponent, Spectral Analysis, Industrial Organisation

JEL Classification: C5, G1, L1

1. Introduction

When modelling price movements, it is common to use a random walk framework. The random walk assumption limits modelling changes in prices over time primarily to using auto regressive and moving average processes. While this technique offers strong short-term forecasting, it cannot offer much of a description about how and why prices are changing over time, aside from correlation to past prices. Since ARIMA (Auto-Regressive Integrated Moving Average) models remove elements of long-term relationships in order to make the data stationary, we lose quantification of the long-run elements in a time series

[1] Assistant Professor Economics/Finance W.H Thompson School of Business, Brescia University, 717 Frederica St., Owensboro KY 42301, rohnn.sanderson@brescia.edu

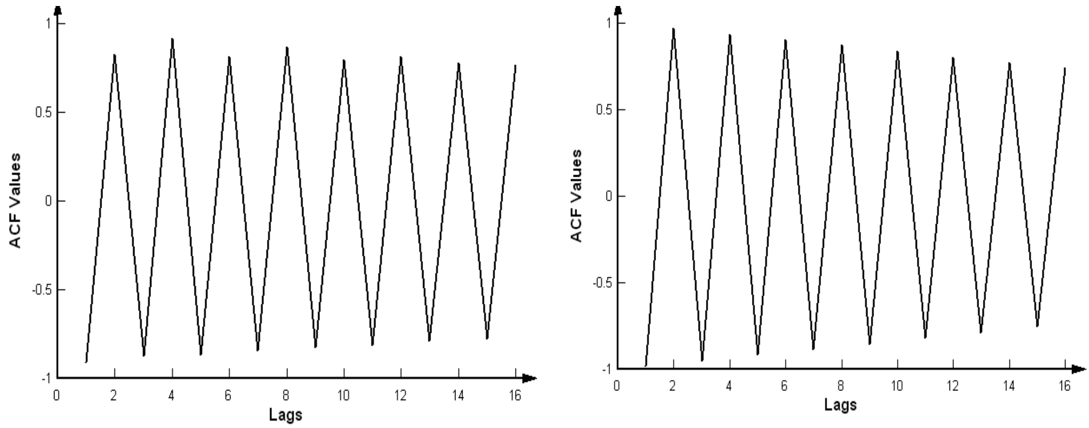
and move those elements into the error term. The complexity of modelling long-term relationships is compounded by the fact that if a dynamic economic system is non-linear, the system may be complex enough to pass standard tests of randomness and become misidentified (Baumol and Benhabib, 1989). The misidentification is usually due to the fact that deterministic processes have infinite variance and we tend to remove infinite variance from data, most commonly in the form of differencing. This critique is not novel, as many researchers have been concerned with measuring longer stationary cycles in dynamic economic systems (Baumol and Benhabib, 1989; Fama and French, 1988; Hsieh, 1991; Lo, 1991; Mayfield and Mizrach, 1992). One determinant of price path behaviour is long run dependence (i.e. memory/history). Highly persistent behaviour in economic systems can lead to events farther back in the history of the series that continue to have an influence on today's prices. The persistence of a series can be measured through the calculation of a Hurst Exponent (Hurst, 1951). While the Hurst Exponent does not definitively determine whether or not a system is linear or non-linear it does aid in our understanding of how the series will propagate in the future.

Gold, not unlike other financial instruments, is subject to memory (long-term cycles). Memory in a financial process implies that the history of prices partially dictates how prices fluctuate in the future. The presence of memory can mean that deterministic behaviour is present in a system. Deterministic behaviour is typically multiplicative in nature. Again this idea is not new to the literature, as many systems have been tested for deterministic and chaotic behaviour. Deterministic behaviour is just one component of the workings of gold prices, there is also a random component as well. This article will take advantage of the Hurst Exponent and a space-time regression in order to separate the deterministic from the random component of the changes in price. Throughout this article, deterministic components may be referred to as endogenous and random components may be referred to as exogenous. This is because the deterministic component represents intra-industry changes that have an effect on the market price, whereas the random component represents the effect on the market price from outside influences. Separation of the two components of the market price can allow us to see the extent to which market structure and macroeconomic changes affect price.

2. Analysis of Components of Gold Prices

Before we start with any data analysis let us get to the root of the long term memory problem. In Figure 1 there are ACF's for two different series: one is a deterministic chaotic logistic equation, the other is Regular Brownian Motion.

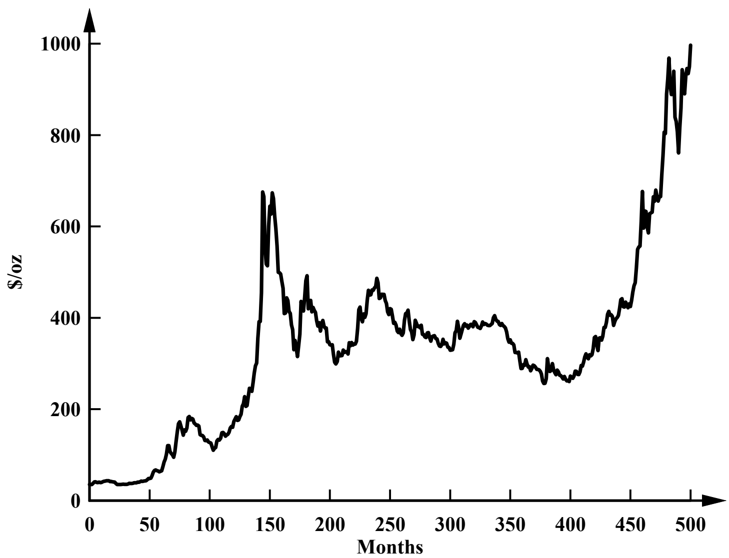
Figure 1: Comparison of a Deterministic and Random ACF



The first panel is the ACF for the chaotic logistic data and the second is the Regular Brownian Motion. Although there are slight differences, there is relatively little that distinguishes the two. In modelling we may see poor performance from an ARIMA model with chaotic data, but that is all. It should be kept in mind that there are many functional forms that can produce chaotic behaviour aside from a relatively simple logistic function. That is the motivation behind looking at the series in its entirety for memory prior to any modification.

The first dataset used in this study is the average monthly gold price per ounce from January of 1968 to October of 2009. The data are displayed in Figure 2 below.

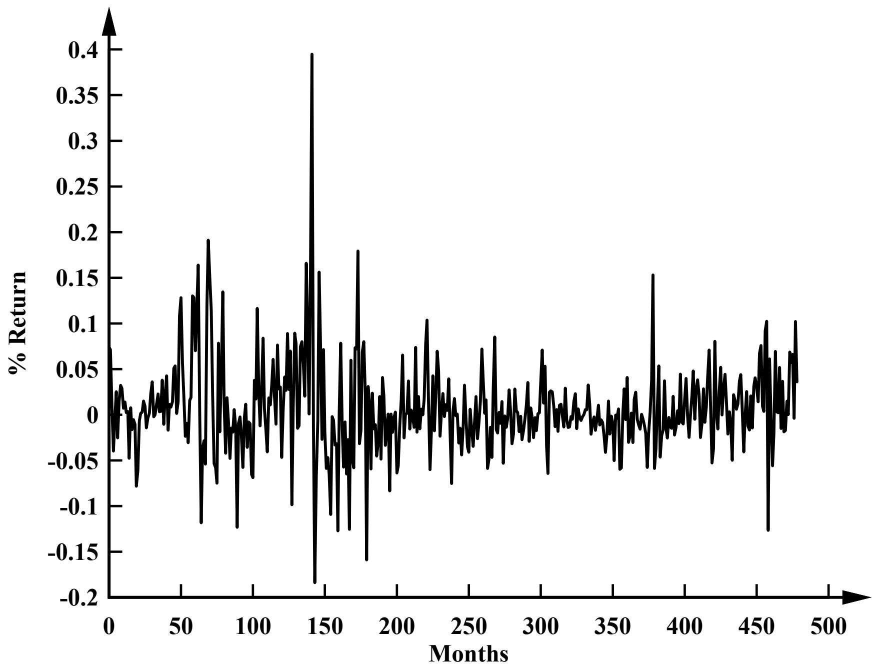
Figure 2: Average Price of Gold by Month (1968-2009)



Source: (Kitco, 2009)

As we would expect, there is a trend and some randomness in the series. Removing the trend by looking at the data of the percentage change in the price shows the stationary form of the data (Figure 3).

Figure 3: Average Monthly Return in Gold Price (1968-2009)



The existence of heteroskedasticity in the error term is also common in many financial series. A test of the squared residuals in both the stationary and non-stationary data shows the presence of heteroskedasticity in the error term. Although we will not do an ARCH model in this article for the sake of brevity, there are most definitely modeling techniques with an ARCH process that could model this behaviour. However we would still not be allowing for the possibility of an infinite variance process. The ACF and PACF plots as well as the non-stationarity of the data would suggest an ARIMA model of order (1,1,1). The results of the ARIMA(1,1,1) in Table 1 are as follows:

Table 1

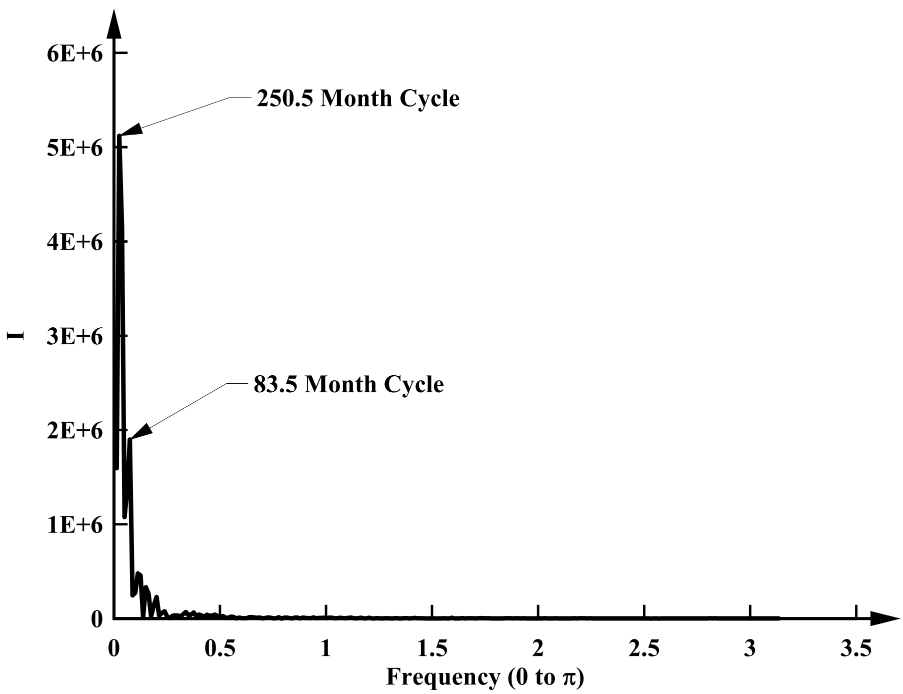
Variable	Coefficient	Std. Error	t-statistic	Significance
Constant	2.6864	1.5295	1.7564	[0.0797]
AR(1)	-.4718	0.1058	-4.4599	[0.0000]
MA(1)	0.7314	0.0834	8.7732	[0.0000[
R^2 = 98.63%	SE = 19.23			

Using a modelling approach such as ARIMA does not measure or investigate the existence of a deterministic long-run cycle with infinite variance. It should also be noted that the Hurst Exponent that we will use to estimate the amount of memory in the series is linked to the differencing parameter in an ARIFMIA model where the Hurst Exponent is equal to 1-d. Instead, we will start our analysis with the reverse question; is there memory or long-run cycles in the data?

To discover long-run cycles we want to impose as little of a functional form as possible and avoid averaging, differencing and the like. Although there are many directions that can be taken to accomplish this we will use a spectral analysis to test for the existence of long-run cycles, due to the acceptance of the technique (Clegg, 2005; Clegg, 2006; Sarker, 2007; Smith, 1992; Stone, Lewi, Landon and May, 1996). Then, the persistence of the memory in the system will be measured via the Hurst Exponent. This will provide a quantification of the level of the long-run effects.

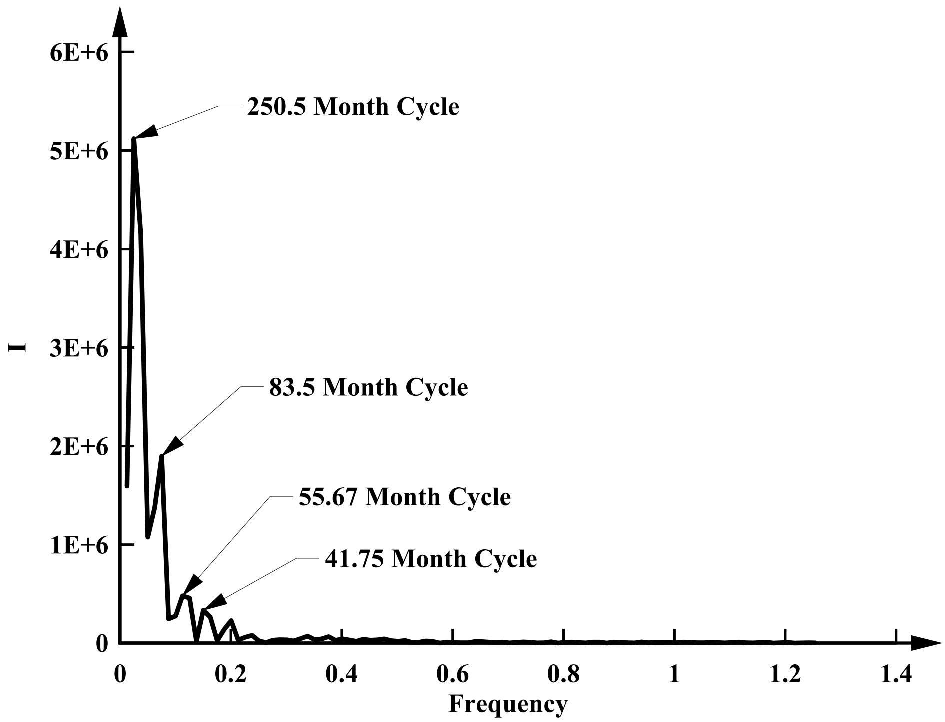
To determine if periodic components exist, a traditional spectral analysis will be used. We will not use a stationary series for this analysis because we want to allow for the possibility of infinite variance in the deterministic process. Instead, we will separate the deterministic and random components by their long term memory and their linear seperability via the ACF. In Figure 4 the only significant cycles at a 5% level are at 250.5 and 83.5 months. Over the rest of the frequency domain the periodicity falls off. The results suggest that the cycles in gold prices occur over very long intervals in time.

Figure 4: Full Spectrum Periodogram for Gold Price



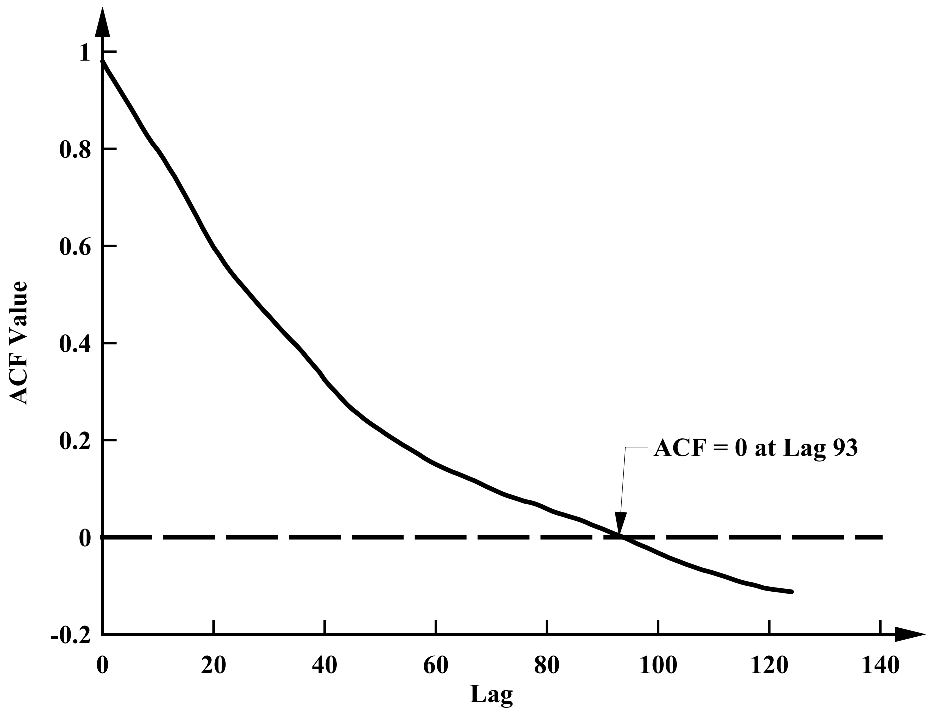
There are two more cycles that are significant at the 5% level that cannot be seen over the entire range of the frequency domain. To see all of the statistically significant cycles, the frequency window has been shortened to 1.4 (Figure 5). In addition to the previously mentioned cycles, there is also a 55.67 and a 41.75 month cycle that are significant at the 5 % level. These additional two cycles are still rather long and there is no statistically significant cycle under three years in length.

Figure 5: Shortened Periodogram for Gold Price



This demonstrates some of the problems with identifying patterns in financial data. Here we have a series that exhibits long cycles over time, which may suggest a certain amount of memory and deterministic behaviour in the system. If ARIMA modelling is used, these long-run cycles will be removed and we will not have a chance at identifying the potential for a portion of the series to have infinite variance. To identify the break between random and deterministic components in a linear fashion, we can measure the dependence through the autocorrelation function. Figure 6 displays the level of autocorrelation within the system, the ACF value does not reach zero until a lag of 93 months. This is where we will separate the data in the space-time regression by the components that have finite variance and the components that do not.

Figure 6: ACF of Gold Price Data



To measure the level of persistence in the system, the Hurst Exponent will be estimated via the periodogram method (this is to keep continuity with the previous spectral analysis). The results are shown in Table 2 and Figure 7 below. The Hurst Exponent is calculated from the regression equation results in Table 1 below.

Table 2

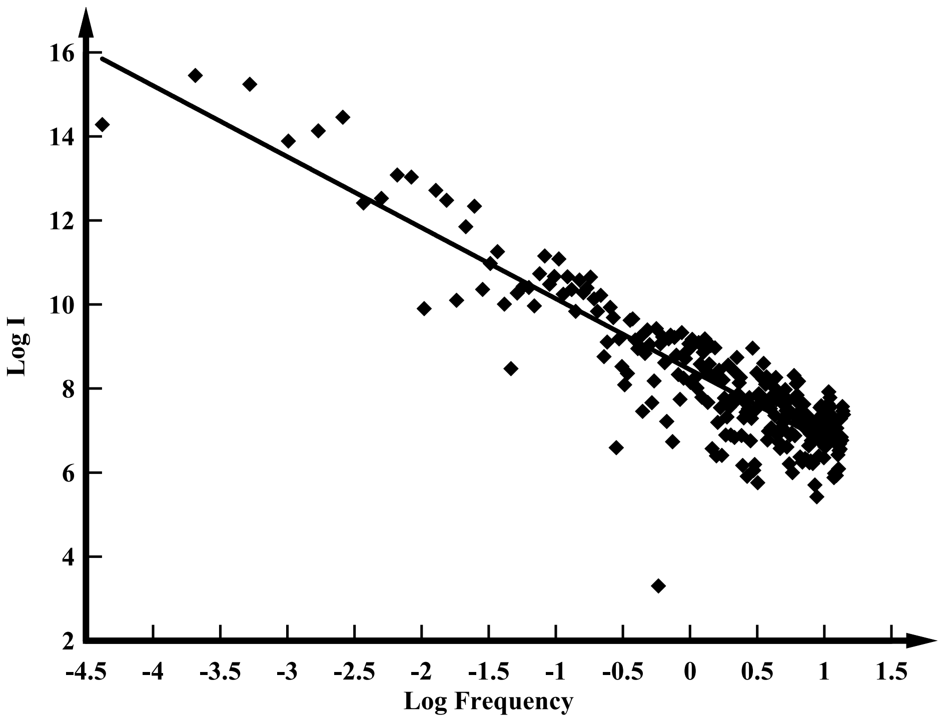
Variable	Coefficient	Std. Error	t-statistic	Significance
Constant	8.448	0.053	160.694	0.000
Log Frequency (α)	-1.689	0.054	-31.229	0.000
R^2 = 79.73%	SE = 0.82			

The Estimate of the Hurst Exponent (H) is:

$$H = {(1-a)}/{2} \quad (1)$$

In this case, the point estimate of the Hurst Exponent is H= 1.345. Given a 95% confidence interval the Hurst Exponent has a range from 1.292 to 1.397. If the system is random (no memory) the Hurst Exponent would be equal to 0.5. This robust result confirms the presence of persistent memory in the system, meaning that history is causing some of the changes in price over time. This suggests that a portion of the structure of gold prices is deterministic.

Figure 7: Hurst Estimation of Periodogram Results

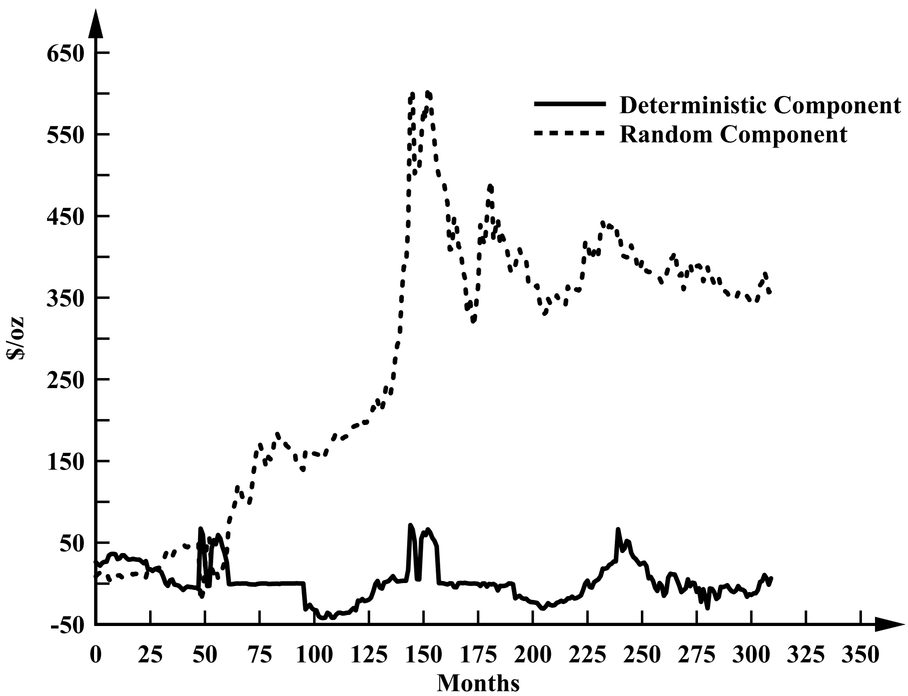


With the information gathered from the spectral analysis, autocorrelation function, and the Hurst Exponent, a space-time regression was performed in order to separate the deterministic from the random components. Since the space-time regression uses the memory information to separate components that are dimensionally independent, we can split the price data into two basic components, which sum to equal the entire signal. It should be retained that deterministic phenomena that are not dimensionally independent may have infinite variance, whereas a random phenomenon does not. Definitionally we will define the two together as the entire price, where:

$$Price_t = Deterministic_t + Random_t \tag{2}$$

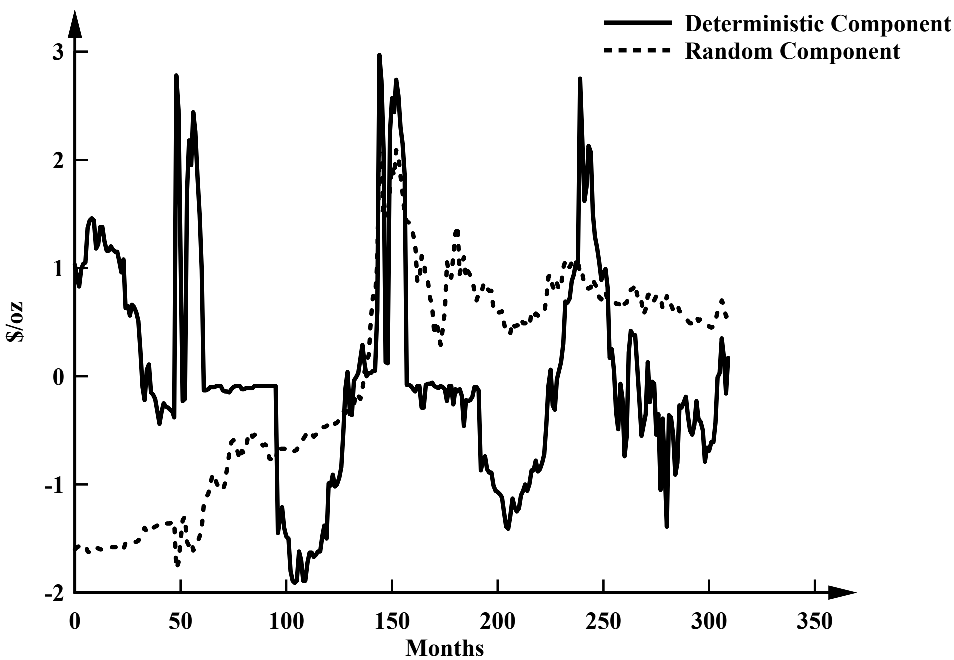
In Figure 8 below the two components of the price of gold can be seen.

Figure 8: Separated Components of Gold Price (1968-1993)



From Figure 8, it can be seen that the deterministic component is the smaller of the two components of gold price. The random (additively separable) component is the larger of the two components. This infers that most of the market price of gold is coming from exogenous events (outside of the gold industry) and that very little of the price of gold is determined by endogenous events. Thus the data suggests that the market structure of the gold industry has little impact on the market price for gold.

Figure 9: Normalized Separated Components of Gold Price (1968-1993)

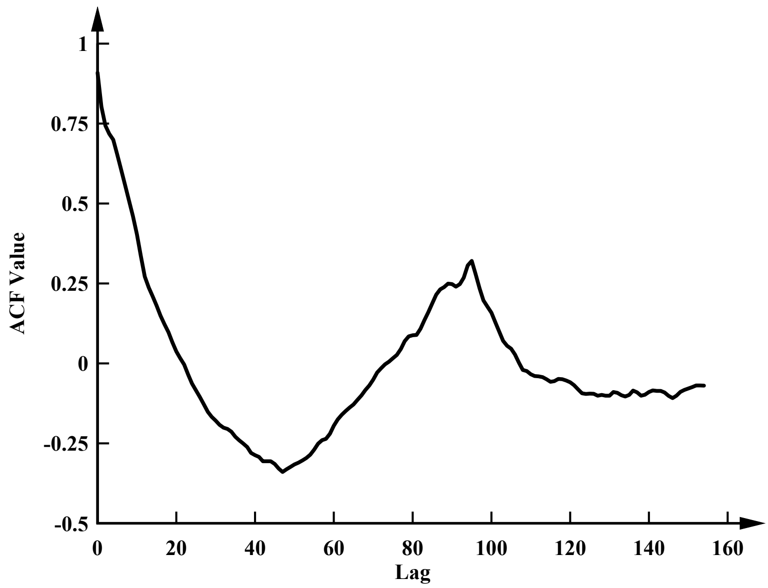


To better identify the behaviour of the two components, both series were normalized (Figure 9) to reduce the effects of scaling.

In Figure 9 it can be seen that the deterministic portion has cycling behaviour and that the additively separable component appears as if it is AR(1). Further investigation of these two components on an individual basis is necessary to determine their effect on market structure. That will not be done here as it is not the focus of this article.

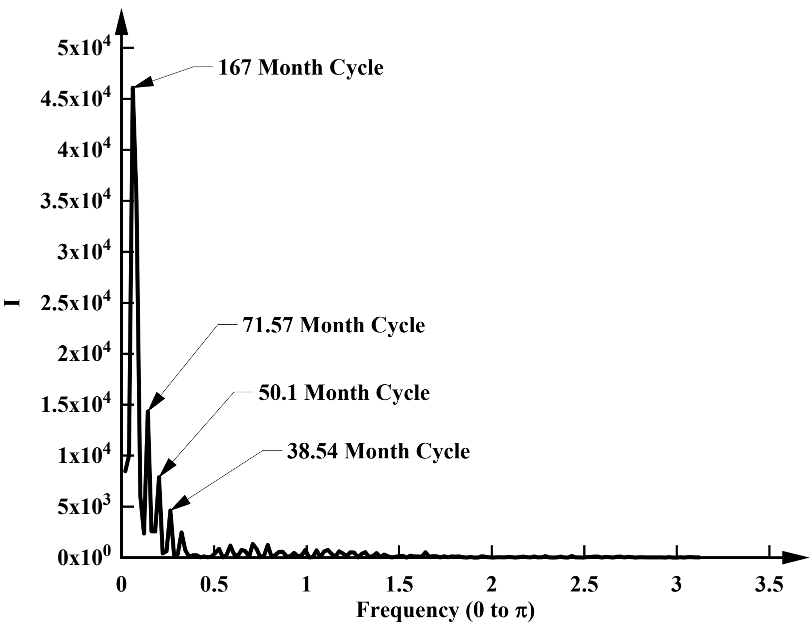
In Figure 10 below, the ACF plot of the deterministic component shows the cycling behaviour. It is important to note that we can now better identify the deterministic component.

Figure 10: ACF Plot of Deterministic Component



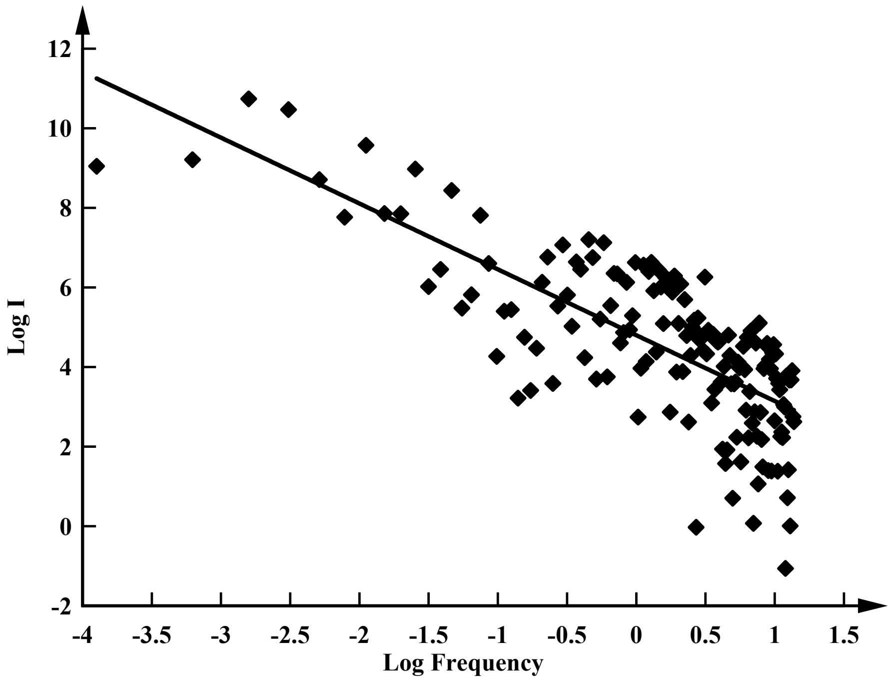
Performing a spectral analysis again on the deterministic component gives us the cycling of the deterministic behaviour (Figure 11).

Figure 11: Full Spectrum Periodogram of Deterministic Component



Now we can see four cycles that are significant at 167, 71.75, 50.1 and 38.54 months. This demonstrates that there is still long-run dependence in the system which is confirmed by another test of the Hurst Exponent (Figure 12).

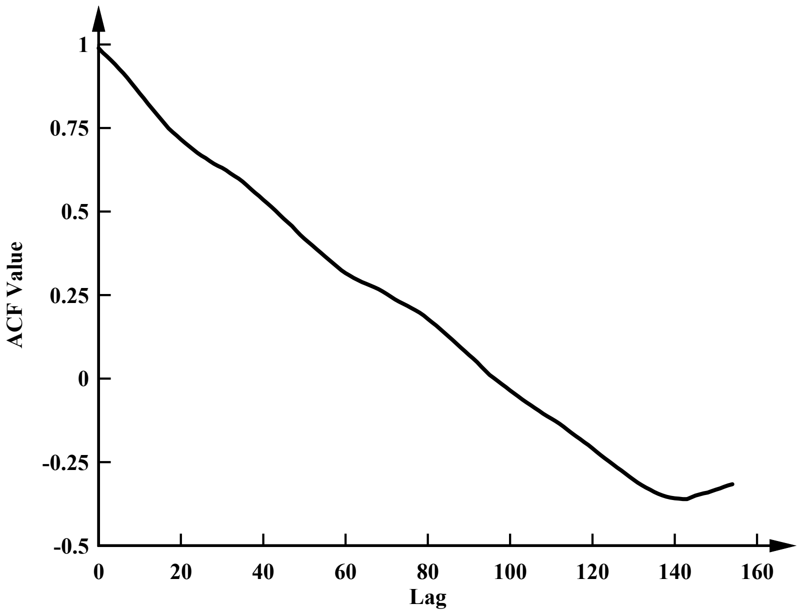
Figure 12: Hurst Estimation of Deterministic Periodogram Results



The Hurst Exponent is equal to 1.509 with a range of 1.414 to 1.604 at a 95% confidence interval, demonstrating that there is memory in the system. One conclusion that it might be drawn is that since there is memory in the endogenous portion of the gold price this displays that the industry itself is not perfectly competitive. If it was perfectly competitive, the endogenous component would attenuate to a flat signal over time. It can be concluded that although the endogenous component is small, market structure does play a role in the market price. However, as it was seen in Figure 8, the market structure impact is minimal in this case. The general result shows that changes in the market structure of the gold industry have very long-run impacts and that the market structure impacts have a small effect on the market price. From a theoretic standpoint this makes sense; although there are few sellers, there are many buyers. Therefore, it is the buyers of gold that are causing the large changes in the equilibrium price. In the case of gold prices, the exogenous component of the price has the greatest affect. Investigation of the exogenous component of the price of gold in more detail is necessary, but it is outside of the scope of this article.

In Figure 13 below, the additively separable (exogenous) component shows a series that has some autoregressive components. It should be noticed that this plot looks similar to the ACF plot of the entire series, which again reinforces the large difference in the magnitudes of the two components.

Figure 13: ACF of Random Component

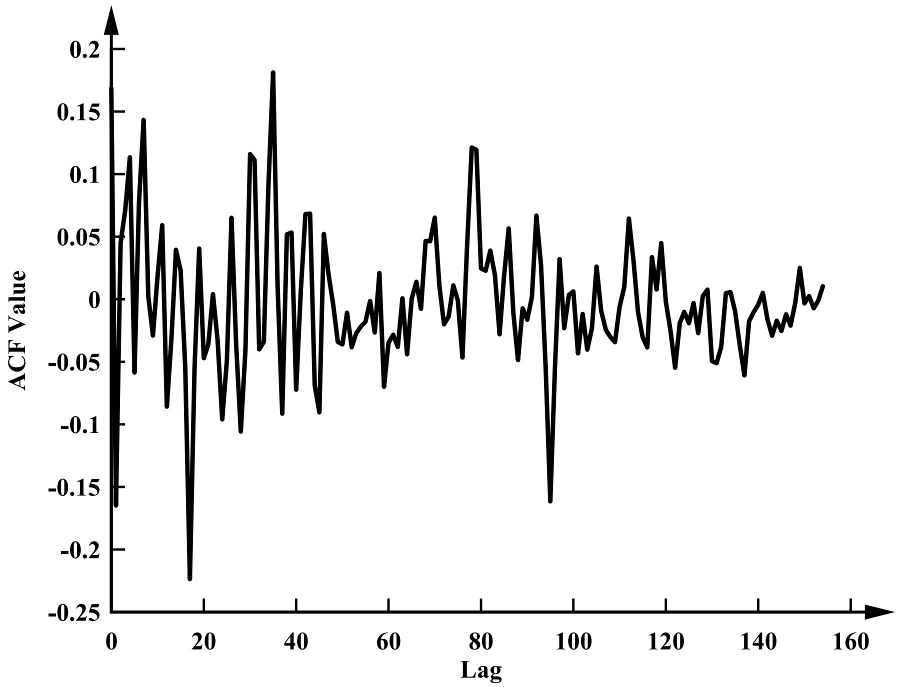


To account for the non-stationarity of the exogenous component, an AR(1) regression with a trend was performed (Table 3). This makes the exogenous data stationary, as it can be seen in Figure 14. Bias has also been removed from the estimators because the endogenous component has been removed.

Table 3: AR(1) Model of Random Component

Variable	Coefficient	Std. Error	t-statistic	Significance
Constant	47.919	78.146	0.613	0.540
Trend	1.279	0.401	3.194	0.002
AR(1)	0.981	0.011	88.121	0.000
R^2 = 98.70%	SE = 18.54			

Figure 14: ACF Plot of Residuals – AR(1) Model with Trend



The AR(1) model is commonly used for financial models of price movements over time. It is true that we could have just modelled the entire price of gold with an AR(1) model and would have obtained similar model results as to those in Table 1. However, the purpose of this technique is not about forecasting per se, it is about being able to compartmentalize prices in a way that helps determine cause and effect. In the case of gold, it was unknown a priori that the endogenous component of the price would be small. It is important to investigate these effects first before a modelling decision is made. For gold prices specifically, we have learned that there are cycles and they are very long. For forecasting purposes this may only be useful for longer time horizons. However, the market structure implications of the result are the most important. The small impact of the market structure tells us the changes in the market structure have little impact on market prices. This type of analysis is important to understand how much market structure changes will impact equilibrium prices. The small size of the endogenous component may not be the case for other commodities or precious metals; each one will need to be tested individually to better understand market structure impacts in those markets.

3. Conclusion

The price of gold has two major components, deterministic and random. In the case of gold prices the deterministic component is small relative to the random component. This suggests that industry structure has little effect on the price of gold. The preponderance of the results of the analysis concludes that external events (randomness) have the largest impact on the changing price of gold over time. This finding may have many important consequences.

For example, from an anti-trust standpoint this type of analysis can give better insights as to how mergers may affect an industry. The case of gold mergers will have very little influence on prices whereas the result may be different in other industries. It is important to note that there could be two industries with the same or similar HHI indices but with drastically different exogenous and endogenous signals that impact their respective markets differently.

From the analysis, we now know that external factors, such as business cycle events, will have a larger effect on price changes than that of intra-industry competition. There is very little that firms in the gold industry can do to alter market prices.

This paper should serve as just the beginning of a process of testing industries along these lines. More research needs to be done with this methodology on other industries to determine if there is true merit to the technique. In the appendix the same analysis is performed on two other commodities for comparison to the results on gold prices. Future extensions with respect to gold prices include determining supply and demand curves for both effects, which was not possible with this data set as the production numbers of gold have been historically unreliable. Further investigation as to what relationship exogenous and endogenous components may or may not have with the HHI and how much they vary with different industries is needed.

In terms of our understanding of long-memory processes, as well as deterministic and chaotic deterministic behaviour is concerned, further research needs to be done in economics and finance to better understand how and if we can use some of the techniques that have been developed in physics and the biological sciences. What we do know is that new dynamical system techniques are being further developed and further investigation of their validity and use in economics and finance is warranted as we continually strive to understand a really complex behaviour.

References

Baumol, W. J., and Benhabib, J., 1989, 'Chaos: significance, mechanism, and economic applications', *The Journal of Economic Perspectives*, 3, 1, pp. 77-105.

Clegg, R. G., 2006, 'A practical guide to measuring the hurst parameter', *International Journal of Simulation: Systems, Science & Technology*, 7, 2, pp. 3-14.

Clegg, R., 2005, 'A Practical Guide to Measuring the Hurst Parameter', UK Performance Engineering Workshop, Newcastle.

Fama, E. F. and French, K. R., 1988, 'Permanent and temporary components of stock prices', *Journal of Political Economy*, 96, pp. 246-273.

Hsieh, D. A., 1991, ‘Chaos and nonlinear dynamics: application to financial markets’, *The Journal of Finance*, 46, 5, pp. 1839-1877.

Hurst, H., 1951, ‘Long-term storage capacity of reservoirs’, *Transactions of the American Society of Civil Engineers*, 116, pp. 770-799.

‘Kitco - Gold Precious Metals - Buy Gold Sell Gold, Silver, Platinum - Charts, Graphs, Prices, Quotes, Gold Stocks, Mining Stocks, bullion dealers.’ Kitco - Gold Precious Metals - Buy Gold Sell Gold, Silver, Platinum - Charts, Graphs, Prices, Quotes, Gold Stocks, Mining Stocks, bullion dealers. N.p., n.d. Web. 18 Nov. 2009. <http://www.kitco.com/>.

Lo, A. W., 1991, ‘Long-term memory in stock market prices’, *Econometrica*, 59, 5, pp. 1279-1313.

Mayfield, E. S. and Mizrach, B., 1992, ‘On determining the dimension of real-time stock-price data’, *Journal of Business & Economic Statistics*, 10, 3, pp. 367-374.

Sarker, M.M.A., 2007, ‘Estimation of the self-similarity parameter in long memory processes’, *Journal of Mechanical Engineering*, 38, pp. 32-37.

Smith, L., 1992, ‘Identification and prediction of low dimensional dynamics’, *Physica D*, 58, pp. 50-76.

Stone, L., Giddy L. and May, R., 1996, ‘Detecting Time’s Arrow: A Method for Identifying Nonlinearity and Deterministic Chaos in Time-Series Data’, Proceedings: Biological Sciences, 263, 1376, pp. 1509-1513.

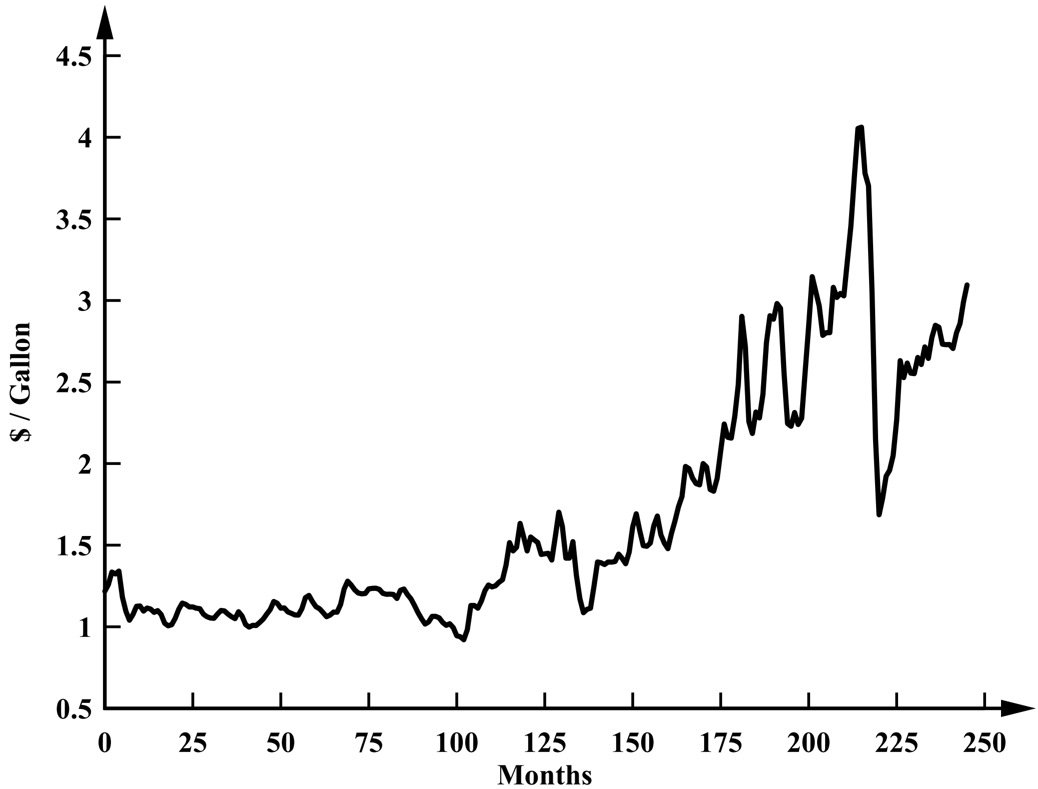
‘U.S. Price of Natural Gas Delivered to Residential Consumers (Dollars per Thousand Cubic Feet)’. U.S Energy Information Administration. N.p., n.d. Web. 21 Feb. 2011. <http://tonto.eia.gov/dnav/ng/hist/n3010us3m.htm>.

‘U.S. Regular All Formulations Retail Gasoline Prices (Dollars per Gallon)’. U.S Energy Information Administration. N.p., n.d. Web. 21 Feb. 2011. <tonto.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMM_EPMR_PTE_NUS_DPG&f=M>.

Appendix

In order to see how this technique works with other data, a similar analysis was performed on the average U.S regular formulation retail gas price from August 1990 to January 2011 (Figure 15).

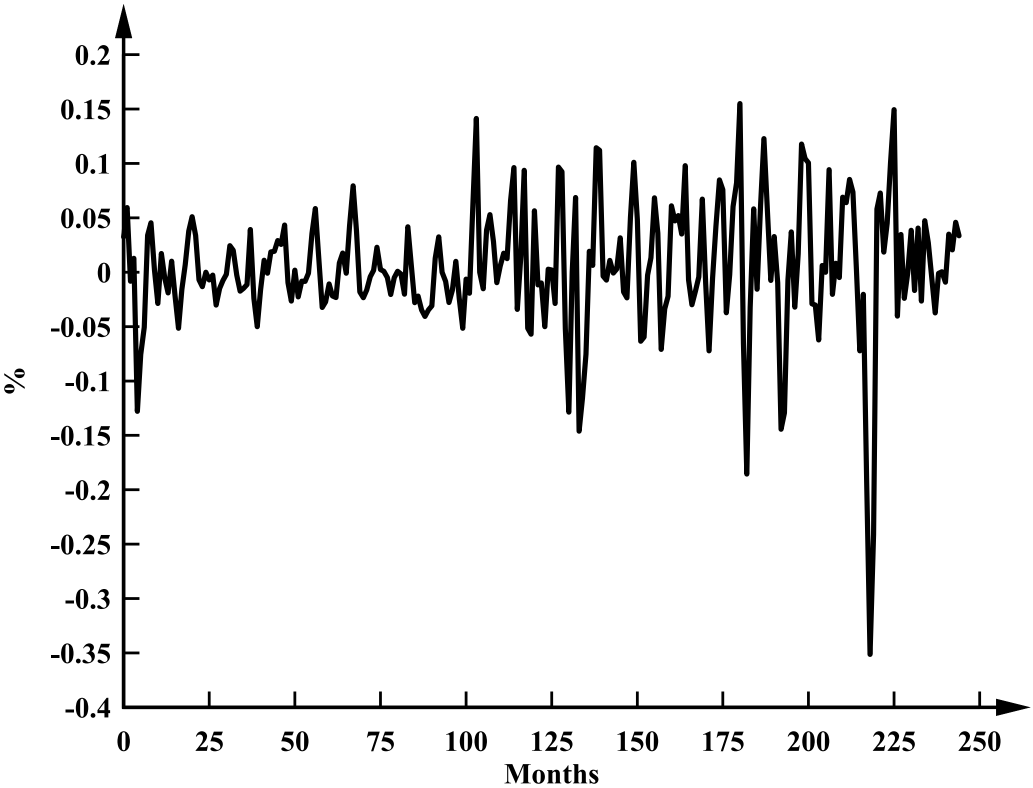
Figure 15: U.S Average Monthly Retail Gas Price (Aug 1990 – Jan 2011)



Source: (U.S Energy Information Administration, Feb-11)

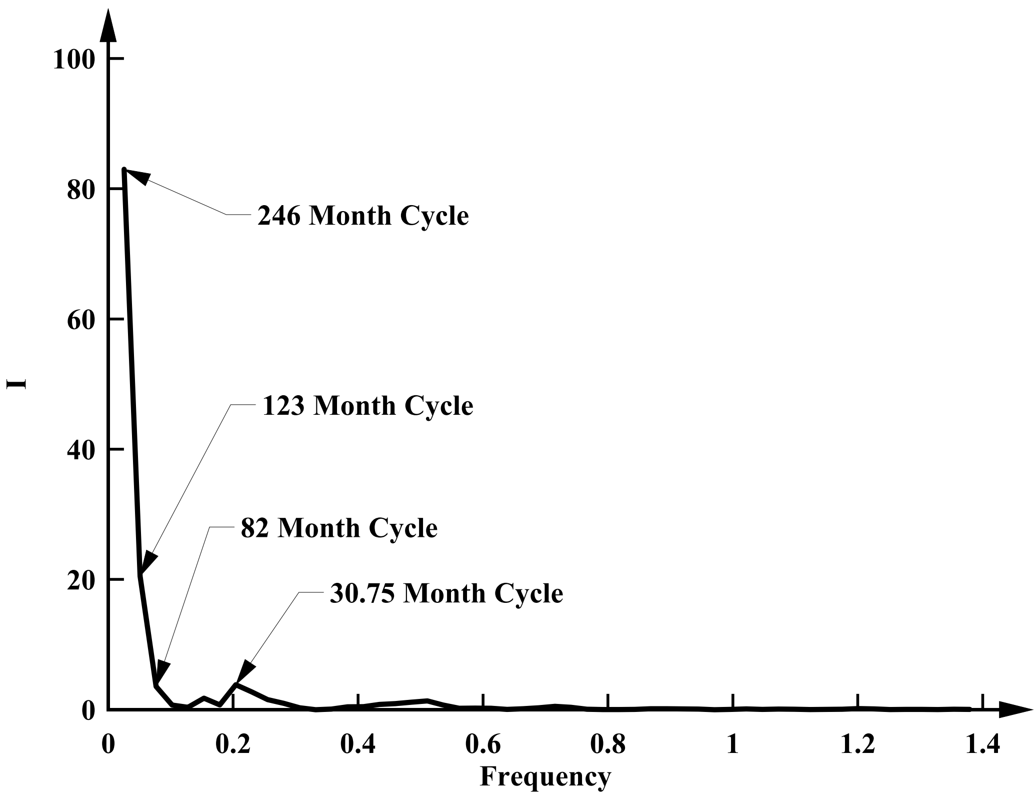
Again we can see similar data analysis problems such as the apparent heteroskedasticity in the data (Figure 16). This can be confirmed through the stationary plot below as well as with a t-test of the squared errors of the series.

Figure 16: Percentage Change in Average Monthly Retail Gas Price



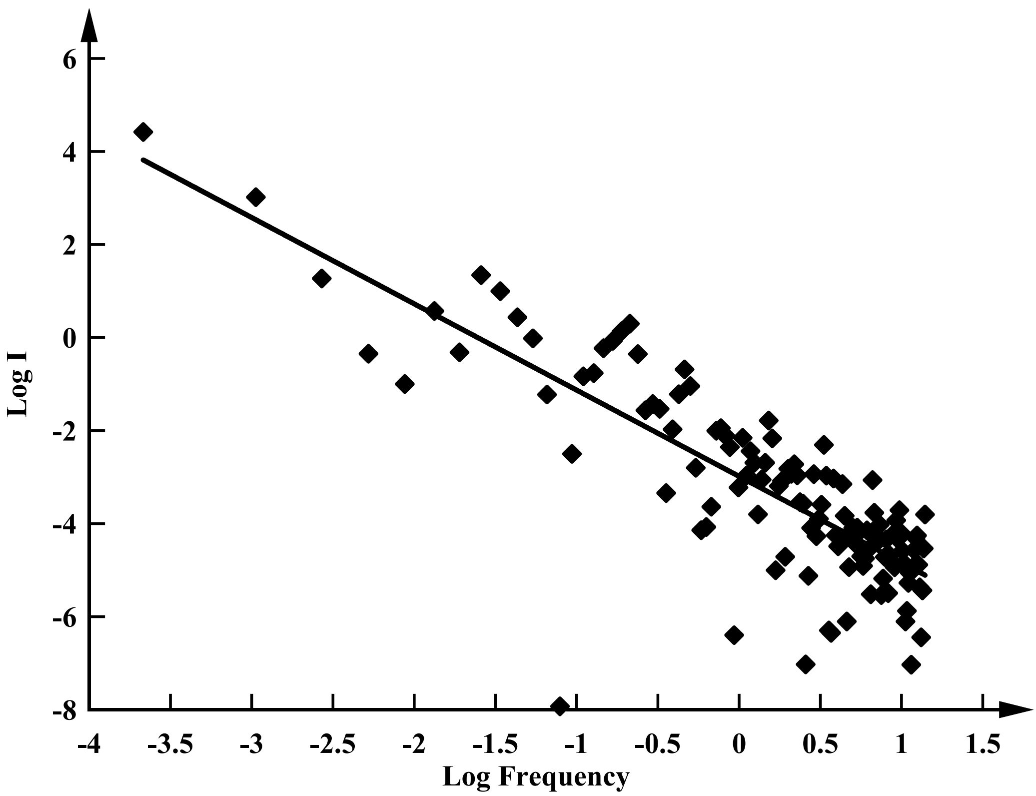
Again let us look at the original series to see what cycles may exist in the data. The following four cycles are significant at the 95% level. Again we see some long-run memory as the shortest cycle is 30.75 months (Figure 17).

Figure 17: Periodogram of Average Retail Gas Price



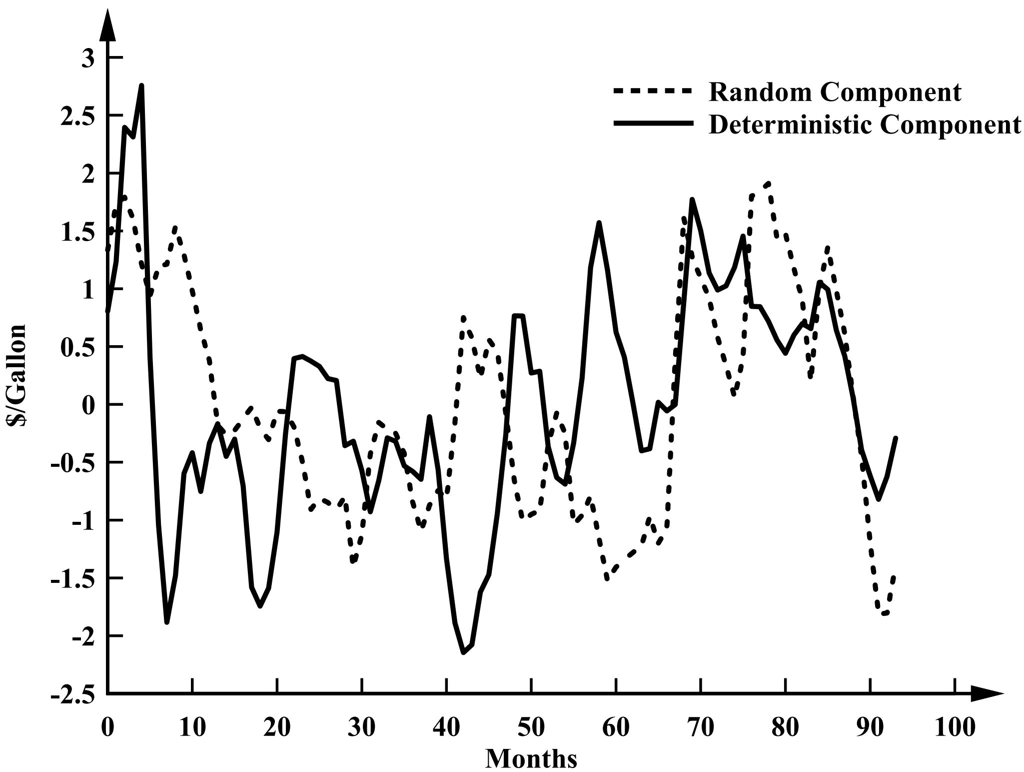
An estimation of the Hurst Exponent confirms that there is memory in the series with the Hurst Exponent being equal to 1.43 (Figure 18).

Figure 18: Hurst Estimation of Retail Gas Price



Performing the same analysis as before, the deterministic and random components of the series are separated and it can be seen below. As in the gold price data the random component is larger than the deterministic component so we will again look at the normalized data (Figure 19).

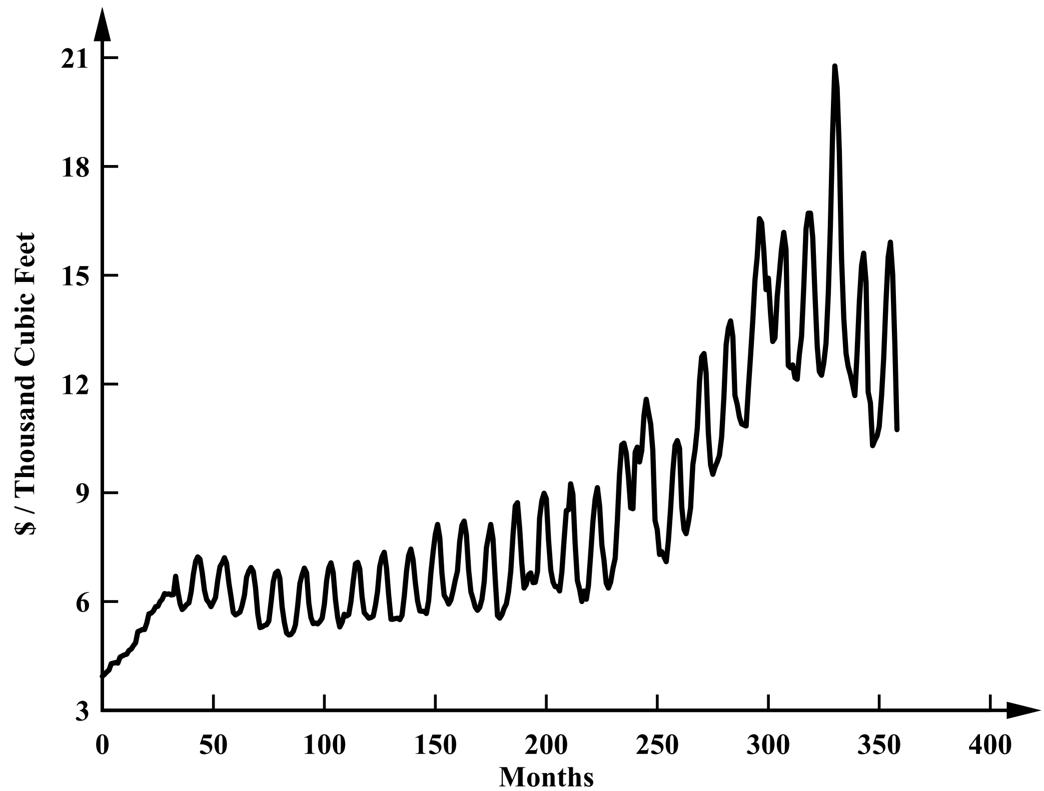
Figure 19: Normalized Separated Components of Gas Prices



In Figure 19, it looks as if both series may be random, again this could be a case where the deterministic component is chaotic. While we will not do so here, we could test the deterministic series for chaotic behaviour with tests as proposed by Stone, Landan and May (Stone et al., 1996). What is more of interest to the author of this article is that we need to allow for its existence when we model behaviour.

Finally, for one more look at methodology we will look at the residential natural gas price in the U.S. as seen in the graph below (Figure 20).

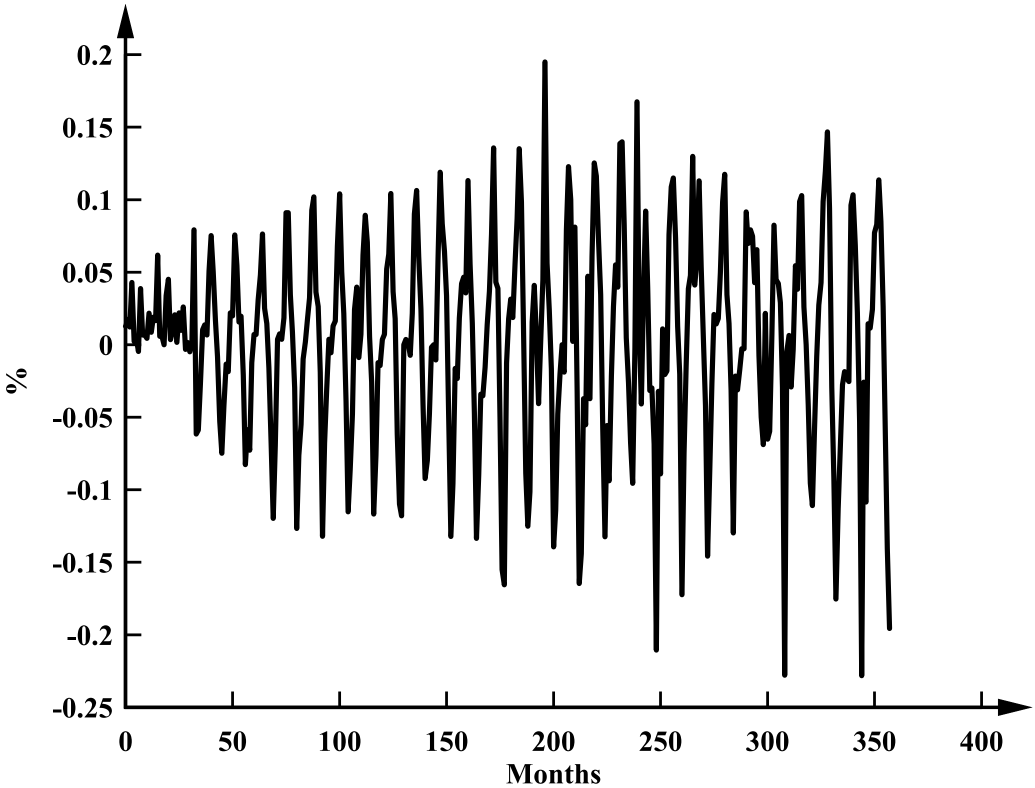
Figure 20: U.S Average Monthly Residential Natural Gas Price (Jan 1981 – Nov 2010)



Source: (U.S Energy Information Administration, Feb-11)

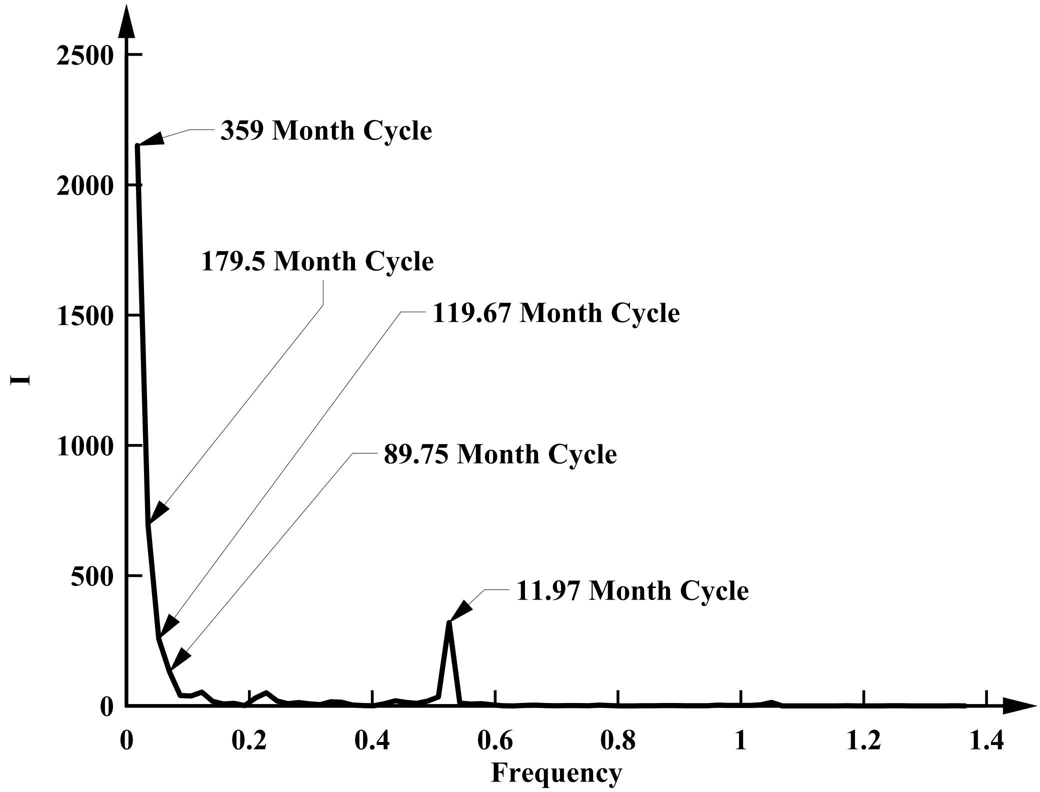
Again we see the same issues with heteroskedasticity, and as before we are confronted with the same methodological issues. Again we could remove the non-stationarity by differencing, but will still have the same methodological issues (Figure 21).

Figure 21: Percentage Change in Average Monthly Natural Gas Price



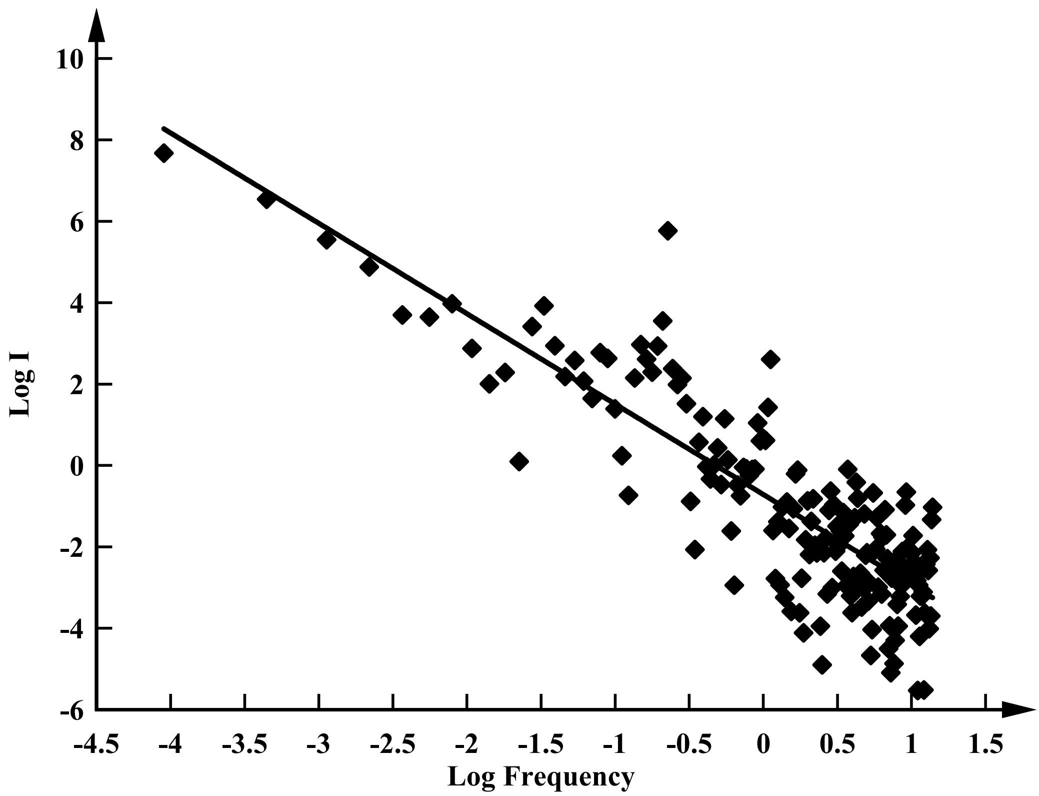
In the case of natural gas, there were five cycles that were significant at the 95% level, the shortest of them being 11.97 months (Figure 22). In this example, natural gas differs from the other two datasets because it does have a shorter cycle, but it is similar in that there is still long-term memory in the series.

Figure 22: Periodogram of Residential Natural Gas Price



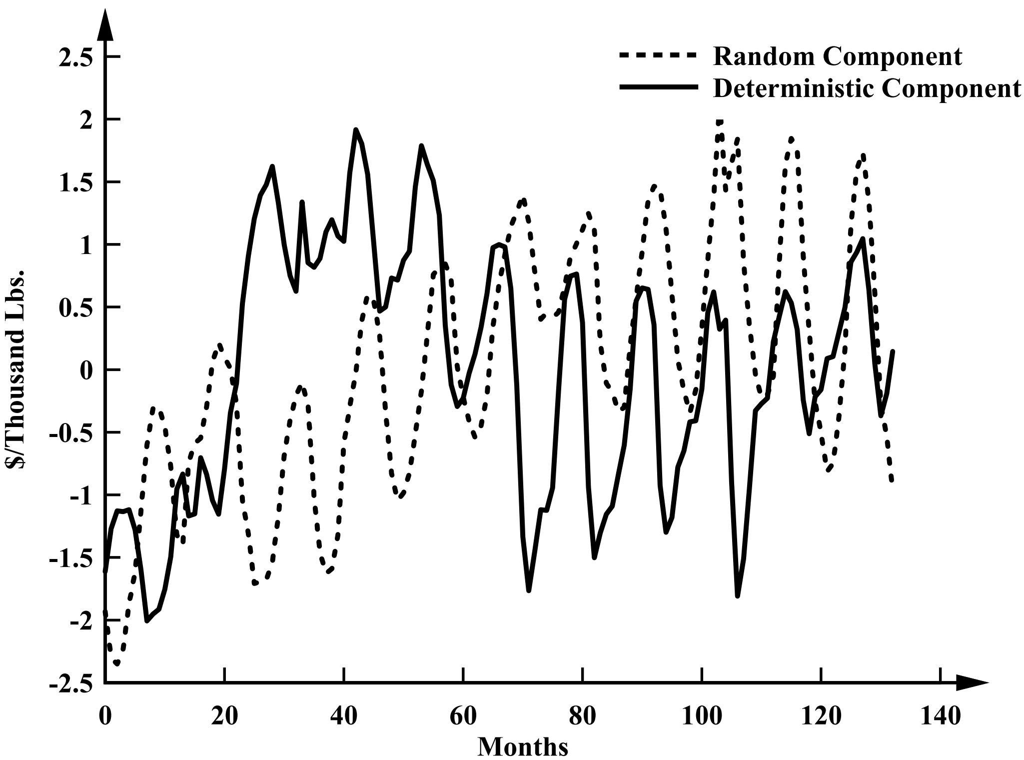
This is confirmed by the Hurst Exponent, which is estimated to be 0.81 (Figure 23), still showing persistence in the data, but at a lower level than the other two datasets.

Figure 23: Hurst Estimation of Residential Natural Gas Price



Performing the same analysis as before, we again find that the random component is larger than the deterministic component. Looking at the normalized data, we can see the behaviour of the two components. In this case, it appears as if the deterministic behaviour has a bit more regular cycling (Figure 24). This could be partially attributed to the lower level of persistence as measured by the Hurst Exponent. What we can see is that the deterministic behaviour in this series has been cycling on a more "regular" frequency than that of the other two series.

Figure 24: Normalized Separated Components of Natural Gas Price



Between all three datasets we can see some similarities and some differences. Why is the random component the largest in all three series? That is a good question that needs to be answered. We also need to ask the question of how prevalent is chaotic behaviour as well as how we can better model chaotic behaviour. These are important questions which hopefully will be answered with future research. A clear point is that we must first start by allowing for the existence of modelling deterministic infinite variance processes and possibly chaotic deterministic processes in order to discover if they are valid or not.