Aalto University
School of Business

Forecasting Finnish Inflation with Commodity Indexes

Economics

Master's thesis

Juhani Koskinen

2014

Department of Economics
Aalto University
School of Business

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Juhani Koskinen

Title of thesis Forecasting Finnish Inflation with Commodity Indexes

Degree Master's Degree

Degree programme Economics

Thesis advisor(s) Pekka Ilmakunnas

Year of approval 2014 **Number of pages** 47 **Language** English

Abstract

We examined whether commodity index prices could improve forecasts of Finnish inflation. Inflation forecasts are made for horizons of one, three, six and twelve months, for the period January 2008 – March 2014. Forecasts from the commodity ADL models are compared to benchmarks set by comparable univariate AR models. In some cases the commodity ADL models have slightly smaller root mean squared forecast errors than their benchmark, but the improvements are not statistically significant.

Data: Finnish inflation, Dow Jones-UBS and IMF commodity indexes, exchange rates

Methods: AR, ADL and MA time series models

Keywords inflation, commodity index, forecasting, time series

Table of Contents

1 Introduction

1.1 Forecasting Inflation with Commodity Indexes

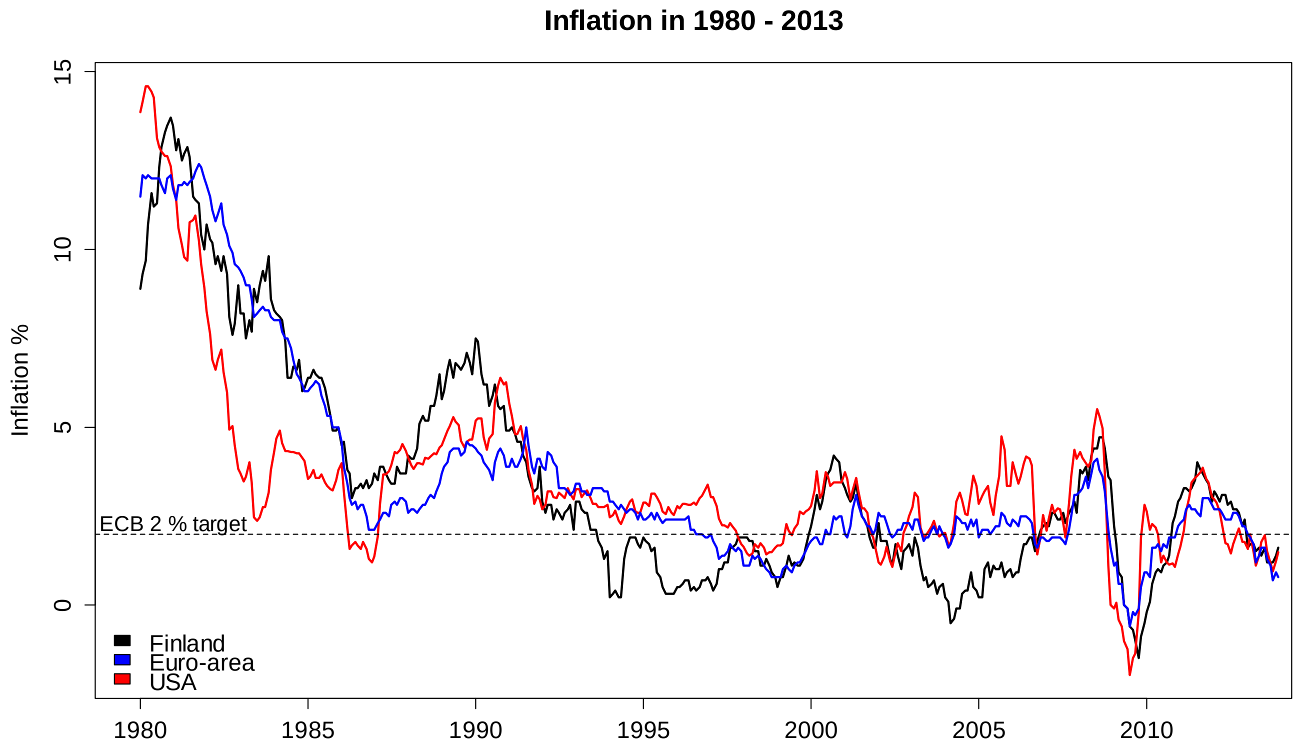
Maintaining price stability by controlling inflation is an important task of central banks. For the European Central Bank, controlling for inflation, by keeping it below, but close to two percent, is their primary objective. The successful management of inflation depends largely on the ability to forecast inflation, so that the central banks can form their policy to reflect also anticipated inflation and not only its current level. Reliable forecasts would allow aiming to cancel the effects from inflationary trends before they are realized. Likewise to driving a car, maintaining course is easier with the ability to see ahead and not only having to rely on the rearview mirror.

Since the mid nineties the level of inflation both in the United States and in the euro-area has been considerably lower than in the 1970s and 80s. The era of double digit inflation already seemed a distant memory. Stock and Watson wrote in 2006 how the decreases in the level and volatility of inflation has led to forecasting becoming easier in the sense of smaller forecasting errors. However at the same time it has become more difficult to find a model capable of improving upon the accuracy of univariate model forecasts. For example a traditional tool in forecasting inflation, the Phillips-curve relation, seems to have broken down, and appears no longer useful in inflation forecasting, as was compellingly argued by Atkeson and Ohanian in 2001.

But the era of "well-behaved" inflation did not turn out to last more than a decade and a half, beginning in the early 90s and ending in the financial crisis of 2008. Since then the level of inflation has sunk even lower and experienced greater volatility as can be seen from Figure 1.[1] The year 2009 saw inflation drop dramatically, but since then the situation has seemed to stabilize. Whether the surge in volatility in 2009 was only temporary or marked the onset of a more unstable inflation era remains to be seen.

[1] Data: Statistics Finland, European Central Bank, and Federal Reserve of Cleveland

Figure 1



Inflation is largely thought to be led by inflation expectations. “Since people’s expectations about inflation influence their behavior in the marketplace, and that, in turn, has consequences for future inflation” (Kwan, 2005). In theory it should be easy to forecast inflation from the expectations, but in practice determining what the actual expectations are has proven difficult. Whilst for example in theory expected inflation should be easy to decipher from interest rates, either from the spread between long and short rates or from a more technical investigation of the yield curve shape, the success in forecasting from these has been limited. Also the explanatory power of the interest rate derived variables is not sustained when past inflation is accounted for (Stock and Watson, 2003).

Inflation appears an autocorrelating progress, in that past inflation “causes”, or provides some information on, future inflation. For a causal process the direction of causality runs forward with the past influencing the future, and by definition the future should have no ability to explain the past. However Ellison et al. (2010) have rather confusingly found that for US inflation the causality runs also backwards which is problematic for the view of inflation as causal.

Forecasting inflation is such a central topic in econometrics that there is a plentitude of literature on almost any kind of model imaginable. Indeed the literature on inflation is so numerous that not even attempts to catalogue it have been attempted in the recent years (Malliaris, 2006).

One possible aid in forecasting inflation is commodity prices. Interest in using commodity prices as a tool for inflation has been in and out fashion, possibly in relation to the levels of commodity prices (Blomberg and Harris, 1995). Commodity prices, and in particular their indexes, could plausible serve as a proxy for inflation expectations. Most prices in a consumer price index are sticky, in that they adjust for inflation only slowly, but commodity prices, both spot and future, are continuously set at auctions so they are adjusted for inflation expectations quickly (Eugeni and Kruger, 1994).

While single commodities, such as gold or crude oil, can potentially serve as leading indicators for inflation, theoretically a bundle of commodities, as found in a commodity index, should provide a more robust indicator. An index is less sensitive to sudden supply shocks from e.g. political situations or bad weather than a single commodity. However the correlation of even seemingly unrelated commodity futures prices is high (Pindyck and Rotemberg, 1990), and recently further more so (Suni and De Meo, 2009) possibly due to commodity-index linked investing (investigated in Stoll and Whaley, 2009), so the benefit gained from diversifying could be diminishing.

Three theories on the linkage between commodity prices and broader inflation were defined by Blomberg and Harris (1995). The first is that commodity prices may give early signals of an inflationary surge in aggregate demand. Illustratively Blomberg and Harris likened the relation between commodity and inflation to the race from the well-known Aesop's tale "The Tortoise and the Hare". "Like the hare, … commodity prices tend to take a quick, early lead in inflation cycles, but ultimately lose to the race, falling in real terms" (Blomberg and Harris, 1995). According to this theory commodity prices first over-adjust to changes in aggregate demand, but the over-adjustment is later reversed. Due to the believed link between commodity prices and economic activity, commodity prices have often been modeled as a function of global economic activity. This demand induced effect on commodity prices should be most pronounced on industrial materials.

Secondly commodity prices and broad inflation should be directly linked, since commodities are an important input into production. In the United States commodities have represented one tenth of the value of production (Blomberg and Harris, 1995). Such direct price effects have historically mostly originated from energy and food commodities. All else being equal the increases in the prices of the commodities end up being passed through to consumer prices.

The third theory is that commodities may be seen as a useful inflation hedge by investors. This behavior would also have self-fulfilling properties, in that the more it is used, the stronger the link would become. Precious metals, especially gold, have been popular as inflation hedges. But it is likely that the importance of this role has diminished, since the proliferation of derivatives has brought more direct ways for inflation hedging, including derivatives on inflation itself.

1.2 Structure of the Thesis

Our aim is to investigate whether commodity indexes are a useful leading indicator in forecasting Finnish inflation as measured by the annual rate of change of the consumer price index.[2] With a series of pseudo out-of-sample forecasts we will investigate whether autoregressive distributed lag (ADL) models of commodity indexes will be able to improve upon the univariate autoregressive (AR) model. In addition a moving average (MA) model and a naive rolling average forecast model comparable to the type used by Atkeson and Ohanian (2001) are estimated. The MA and AO models are included since they have a significant role in inflation literature. The MA model has had excellent success in forecasting recent inflation (Stock and Watson, 2006), and the AO model which is by intentional design very simple, still has been a difficult benchmark to beat (Stock and Watson, 2006 and, Atkeson and Ohanian, 2001).

Following the example of Stock and Watson (2003) forecasting is performed with pseudo out-of-sample forecasts estimated for a 75 month period from January 2008 to March 2014. This means that a series of forecasts will be estimated, using a dataset which recursively grows with the inclusion of the most recent values. Thus for example for forecasting the next months values, the first forecast will be made from the data from February 1992 (the data for January was lost due to differencing) to December 2007, the second estimate will also include the values from January 2008, the third February 2008, and so on until the last forecast for March 2014 which is estimated on all the history up to February 2014.

---

[2] $\pi_t = \Delta^{12} \log CPI_t$

Forecasts will be made for the next 1 month, 3 months, 6 months and 12 months. When the forecast is made for a longer horizons than one month, the most recent values in the dataset are retarded the equivalent amount, thus for the first one year forecast for January 2008, the most recent data would be from January 2007.

The previous literature, which is mostly on forecasting US inflation, has found that on the whole, despite the theoretical plausibility, commodity prices are not a reliable indicator for inflation. They've been found to improve forecasts during some periods, but not in others. Even in the periods when they've been found to improve the forecast, the improvement has been minor and insignificant. Additionally, inconveniently for a leading indicator in some cases the relation has been found negative so that a rise in commodity prices has signaled a decrease in consumer price inflation (Blomberg and Harris, 1995; and Cecchetti et al., 2000).

Our finding is that some of the indexes improve on the pure autoregressive forecast, on some forecast lengths, but the improvements of the ADL on the AR models are not statistically significant. While the improvements on the forecasts are not significant, the regression coefficients provide some insight into Finnish inflation dynamics. At the shorter horizons the changes in commodity prices and inflation mostly have a positive correlation, but this direction is reversed at the longest one year horizon. This can be seen as evidence for a "Tortoise and Hare" relation between commodities and inflation. At all forecast horizons the MA model has the smallest forecast errors. The Atkeson and Ohanian (AO) model performs the worst of all models at each forecast horizon.

This thesis proceeds so that the next second chapter presents the existing literature on the link between commodities and inflation. The third chapter introduces the commodity indexes used in this study, and the fourth chapter details the data used. The models used are presented in the fifth chapter, followed by an examination of the results in chapter six. Finally chapter seven contains the concluding thoughts.

2 Research on Commodities and Inflation

A wealth of research exists on the link between commodity prices and inflation, or more generally expressed consumer prices. The overwhelming majority of the research is on US inflation, but countries like Australia and Canada (Bloch et al., 2006) have also been covered. It seems that no prior research exists in which Finnish inflation has been forecasted by commodity prices. In this chapter all the articles mentioned explore US inflation, unless otherwise stated.

The general finding has been that commodity prices used to be a useful leading indicator for inflation in the high inflation era of the 1970s and early 80s (Blomberg and Harris, 1995). But since the 1980s the situation has changed and at least with linear models successfully forecasting inflation from commodities has become difficult. We have classified examples of the past research into three categories: Successes, Mixed Results, and Failures, and these are listed in Table 1. The classification is not based on the respective authors on views on the success of their results, but rather on our own rubric. The Successes are examples where the commodities are without reservations judged to be useful in inflation forecasts. In the class of Mixed Results, the commodities were found to make marginal but insignificant improvements on inflation forecasts. Finally in the class of Failures the commodities were found to contain no additional information useful in inflation forecasting purposes.

Table 1

Past Research Classified by Forecasting Performance

	Area	Period	Frequency	Model
Successes				
Edelstein (2007)	US	1993-2004	Monthly	Bagging, Bayesian, Shrinkage & Factor
Browne and Cronin (2009)	US	1959-2008	Quarterly	Vector ECM
Gospodinov and Ng (2010)	G7	1983-2008	Monthly	PC's of convenience yields
Mixed Results				
Webb (1988)	US	1954-1988	Monthly	VAR
Moosa (1998)	OECD aggregate	1972-1993	Monthly	Causality tests
Blomberg and Harris (1995)	US	1970-1994	Monthly	VAR
Cecchetti, Chu and Steindel (2000)	US	1975-1998	Quarterly	ADL
Stock and Watson (2003)	G7	1959-1999	Quarterly	ADL
Cecchetti and Moessner (2008)	19 countries	1992-2008	Monthly	ADL
Acharya (2010)	US	1957-2005	Annual	VAR
Failures				
Boughton and Branson (1988)	G7 aggregate	1962-1987	Monthly	Polynomial distributed lag
Eugeni and Kruger (1994)	US	1970-1994	Monthly	ADL
Garner (1995)	US	1983-1994	Monthly	ADL
Furlong and Ingenito (1996)	US	1960-1995	Monthly	VAR
Mahdavi and Zhou (1997)	US	1958-1994	Quarterly	ECM

We will now take a closer look at the research listed in Table 1. First we present which commodity indexes have been used. The most common sources of the commodity prices have been the Commodity

Research Bureau (CRB), Journal of Commerce (JOC) and the International Monetary Fund (IMF). Blomberg and Harris (1995), Eugeni and Kruger (1994), Garner (1995) and Webb (1988) used both the CRB and JOC commodity indexes. In addition Acharya et al. (2010), Browne and Cronin (2009), Cecchetti and Moessner (2008), Furlong and Ingenito (1996), Gospodinov and Ng (2010) used the CRB index. The JOC data was used by Cecchetti et al. (2000) and Mahdavi and Zhou (1997), while the commodity prices collected by the IMF were used by Boughton and Branson (1988), Edelstein (2007) and Moosa (1998). The choice of commodity index does not seem to affect the results too much; none of the used commodity indexes has come out clearly superior to the others. This is not surprising since in any case the correlations between single commodities are high. On the other hand it is possible to improve forecasts by forming an index with customized weights as was done by Edelstein (2007).

Notably all of the successes are fairly recent: Edelstein (2007), Browne and Cronin (2009), and Gospodinov and Ng (2010) who all achieved to demonstrate that commodities can still serve a useful purpose in inflation within the right framework. What separates these from the earlier less successful examples of research are the models used which are more state-of-the-art than the linear ADL and VAR (vector autoregression) models typical of the earlier research. A possible conclusion to be drawn from this is that the most fertile paths for future models on inflation forecasting lie outside the linear specifications of the ADL and VAR models.

The framework of this study is the more traditional linear ADL models. Therefore the results of these recent successes are not directly comparable to this study. Still it is illuminating to briefly outline their achievements. Gospodinov and Ng (2010) used bootstrap inference on the principal components of commodity convenience yields.[3] Edelstein (2007) used various techniques like bootstrap aggregating (or bagging), Bayesian model averaging, shrinkage estimation and factor models. Unlike Gospodinov and Ng (2010), Edelstein (2007) did not transform the commodity prices in any fundamental way, and solely formed customized indexes of them based on the models used. The valuable insight of Browne and Cronin (2009) is that the bivariate relationship of commodity and consumer prices needs to be augmented by the quantity of money. The introduction of a monetary measure could plausibly resolve one of the current dilemmas of frequent negative correlation for commodity and consumer prices.

[3] The convenience yield (CY) at time t with delivery at time $t + n$, and interest i, is defined as:

$$CY_{t,n} = Spot_t(1 + i_{t,n}) - Future_{t,n}$$

For the rest of the research making up the Mixed Results and Failures the types of models used were all basic linear models. Within these two groups the results are fairly uniform, the difference between the groups being that to qualify as a failure the commodity model needed to be forecast worse than the comparable no-indicator model. For qualification as a Mixed Result the model had to outperform the no-indicator model, but with an insignificant and minor amount. The results of our investigation are inline with this prior research, in that they qualify as a Mixed Result.

2.1 Cointegration

An essential question in the dynamics of the inflation and commodity price relation is that of cointegration. Both commodity prices and inflation are commonly considered to contain one unit root and the question of cointegration is to determine whether they share a common unit root. The past literature has conflicting results on the matter. The conflicting results in themselves can be seen as evidence against cointegration. The economists desire to discover cointegration, is primarily for the coherence of economic theory, and the marginal forecast improvement from the error-correction term in a cointegration based model has been found insignificant (Mahdavi and Zhou, 1997).

The question of cointegration is not examined in a significant share of the commodity-inflation literature at all. These omissions are possible to count as evidence for viewing cointegration unlikely as well. Boughton and Branson (1988) examined the inflations of G7 countries and did not find evidence for cointegration. Likewise Furlong and Ingenito (1996), Moosa (1998), and Ciner (2011) who all tested the US inflation did not find support for cointegration.

In the cases where evidence for cointegration has been found, the type of cointegrating relation identified has not always been the same. Blomberg and Harris (1995) found cointegration between the levels of monthly US inflation and commodity prices. Conflictingly Mahdavi and Zhou (1997) found cointegration between the levels of the consumer and commodity price indexes over the period 1958 – 1994 at the quarterly frequency. However they did not find cointegration between the price of gold and the consumer price index in the period 1970 to 1994. Kyrtsou and Labys (2006) found cointegration between the monthly levels of consumer and commodity price indexes. While they find the cointegration linear, they state that the short-run dependency is nonlinear and chaotic.

Finally Browne and Cronin (2010) find cointegration so that both the consumer and commodity price indexes are each in turn cointegrated both with output, GDP, and the M2 variable for nominal money. Their framework is related on the "price puzzle" as discussed by Hanson (2004) on how in a counterintuitive fashion contractionary monetary policy by the Federal Reserve has led to a rise in the level of consumer prices.

The possible cointegration between the levels of Finnish inflation and the commodity indexes will be tested for in this thesis. Since we are investigating inflation, the possible cointegration between the levels of commodity and consumer prices is left outside the scope of our work. Likewise the further permutations of possible cointegrating relations between the prices and money or output will be left for further research. Also it must be noted that investigating the role of money for Finland is less trivial than for the United States, because the regime change from the Bank of Finland to the European Central Bank inconveniently bisects the history for which the commodity indexes are available (from the 90s onwards).

3 Commodity Indexes

The commodity indexes used here are the non-fuel and with fuel versions of the IMF commodity index and three different versions of the Dow Jones-UBS (DJUBS) commodity index, which are freely available from the IMF and Dow Jones Indexes respectively. The IMF index is of spot prices and the Dow Jones indexes are formed from futures prices.

Besides the DJUBS, another widely followed commodity futures index is the Goldman Sachs Commodity Index (GSCI).[4] An important reason why these two are so followed is that numerous exchange-traded funds (ETF) and other investment vehicles are based on tracking them. They both consist of futures contracts expiring in either one or two months, their main difference being that the GSCI has a much heavier weighting in energy and crude oil in particular. This is because the GSCI is weighted by the world production quantities, while the DJUBS weighting is capped so that no single commodity may constitute over 15 % (or including the products derived from it 25 %) of the index (Dow Jones Indexes, 2010). Unlike for the DJUBS, the GSCI historical data appears not to be freely

[4] http://www.goldmansachs.com/what-we-do/securities/products-and-business-groups/products/gsci/

available, so it could not be included in this study. However since the IMF commodity index is also production weighted, it surrogates well for the GSCI, albeit with spot instead of futures prices.

There are numerous specifications for the DJUBS commodity index, between total return and excess return, currency specified in and type of contracts used. The indexes are composed of futures contracts and 3 Month US Treasury Bills (Dow Jones Indexes, 2010), the total return index is the return from this composite, including the returns from the rolling of futures, interest derived and from the change in interest rates (which is not much, when the duration is short), and the excess return is the total return after the removal of the return effects from the interest rate. In this study only the excess return indexes are considered, while the interest rate element included in the total return index could also plausibly aid in forecasting inflation, it would unnecessarily confound the relationship between inflation and commodity prices. For similar reasons the commodity prices are transformed to euros (or to Finnish markka's (FIM) prior to 1999). The US dollar / euro exchange (USD / EUR) rate could also serve some role in improving the forecasts since declining exchange rates can be a signal of increasing inflation (Cecchetti et al., 2000), so it is not a foregone conclusion which of the series, the one denominated in US dollars or the one in euros, forecasts better.

Of the excess return series for DJUBS the series chosen are the "standard" series of DJUBS consisting of futures expiring in the next one or two months, the DJUBSSP consisting of theoretical spot prices and the DJUBS3M which is a version of the DJUBS with futures expiring in three months. Despite its name the DJUBSSP is also derived from futures and not actual spot prices. However the effects on return from the rolling of futures contracts have been cleaned away from it so that it more precisely provides a general estimate of the trend in commodity prices. The details on the formulation on the Dow Jones indexes can be found in The Dow Jones-UBS Commodity Index Handbook (Dow Jones Indexes, 2010).

The standard DJUBS index is available from Dow Jones also in euro-denominated form, but the normal USD denominated series were chosen for each three versions. This was both due to consistency, so that all the commodity indexes are transformed with the same exchange rates, but also to allow the pre-euro era (until the end of 1998) data to be transformed to reflect the FIM / USD exchange rate.

The International Monetary Funds (IMF) index is a production weighted index of spot prices.[5] As in the case with DJUBS numerous versions of the IMF index are available, differing on which of the product groups are included. The index chosen is the most expansive index with all commodities included. There are two versions of this all commodity index one with energy excluded and another with energy included. Both of these are used in this study.

Thus in total five different indexes, the DJUBS, DJUBS spot (DJUBSSP), the DJUBS three month forward (DJUBS3M), the IMF all commodity including energy (IMF), and the IMF all commodities excluding energy (IMF No Fuel = IMFNF) are chosen. The composition of the indexes is presented in Appendix 1. The indexes cover spot and futures priced, as well as production weighted, and the more evenly balanced DJUBS weighting, and an index without energy. Nevertheless the correlations of the indexes are high as can be seen from Table 2.[6] It is possible that comparing the forecasting performance of the different indexes will shed some light on Finnish inflation dynamics, especially in regard to comparing the IMF and IMFNF to see how much of an effect is from energy prices.

Table 2

Commodity Index Correlations

	DJUBS	DJUBSSP	DJUBS3M	IMF	IMFNF
DJUBS	1.000				
DJUBSSP	0.950	1.000			
DJUBS3M	0.960	0.996	1.000		
IMF	0.951	0.986	0.985	1.000	
IMFNF	0.876	0.846	0.847	0.895	1.000

4 Data

The data for Finnish inflation was acquired from Statistics Finland.[7] The series chosen was the monthly series for the annual rate of change in the consumer price index. The consumer price index is calculated from 50,000 different prices on 486 consumer products and services collected from 2,700 shops, and the prices are always collected in the middle of the month. In addition about a thousand prices are collected in a centralized manner. The consumer price index consists mainly of products and services provided by the private and public sectors, and from value added tax and direct taxes levied on

[5] http://www.imf.org/external/np/res/commod/index.aspx
[6] Monthly series correlation of euro denominated levels. DJUBS series aggregated with monthly medians.
[7] http://stat.fi/til/khi/2013/12/khi_2013_12_2014-01-14_tie_001_fi.html

commodities. The taxes represent roughly a quarter of the index.

Data for the IMF commodity indexes was collected from the IMF and for the three different DJUBS indexes from Dow Jones Indexes. The Dow Jones series are of daily frequency and available beginning from January 1992. The IMF series are monthly and IMFNF is available from January 1991 and the regular IMF index from January 1992. For overall consistency the same sample period is used for all the models, so the first available year for IMFNF is not used and all commodity series used begin January 1992.

Since all of the indexes are denominated in US dollars they ought to be converted to euros. As stated by Boughton and Branson (1988): “In order to isolate the effects of commodity price movements on inflation from those of exchange rates, it is desirable … that commodity and consumer prices be denominated in the same currency.” For dates up to the end of 1998 the indexes are converted to euros through the Finnish mark and the irrevocable exchange rate of one mark equaling 5.94573 euros. The exchange rate data for FIM / USD (up to December 1998) is from the Bank of Finland (BOF) and the EUR / USD rates are from the European Central Bank (ECB).

For the IMF series the conversion to euros is straightforward. Since they are monthly series, the most sensible way to convert them should also involve some type of monthly frequency exchange rates. Therefore they were converted with monthly average exchange rates. Starting from January 1999 monthly average exchange rates are directly available from ECB. For earlier dates the average monthly exchange rates were computed from the daily FIM / USD exchange rates.

4.1 Transforming from Daily to Monthly Frequency

The Dow Jones series are of daily frequency, so they have to be changed into monthly frequency. A common way to summarize the info is to simply use the last value from each month. Since daily futures price series are typically very noisy, Armesto and Gavin (2005) suggest reducing some of this noise by using the monthly means or medians is advisable for forecasting purposes.

First the daily Dow Jones series were converted to euros by using the daily exchange rates from ECB and BOF. Since the DJUBS has a value only for all American business days, and the exchange rates from ECB and BOF have values only for European (and Finnish) business days, this causes that some of the dates cannot be joined. As there are only a few of such dates in a year, omitting these dates does not make too much of a difference. Especially since the daily series are aggregated into monthly values, a missing day in some months presents no meaningful loss of information. As an example for the year 2013 the exchange rate was available for 255 days, the DJUBS for 241, and these coincided for 239 dates. Thus only 2 date's worth of the DJUBS index ended up discarded for 2013 and the situation has to be similar for the other years as well.

These euro-dominated series are then summarized to monthly series by either taking the last value of the month or by monthly means and medians. Their volatilities, or standard deviations, during the years 1992-2007 are reported in Table 3.

Table 3

Monthly Volatilities (And Reduction Compared to Last Day Value)

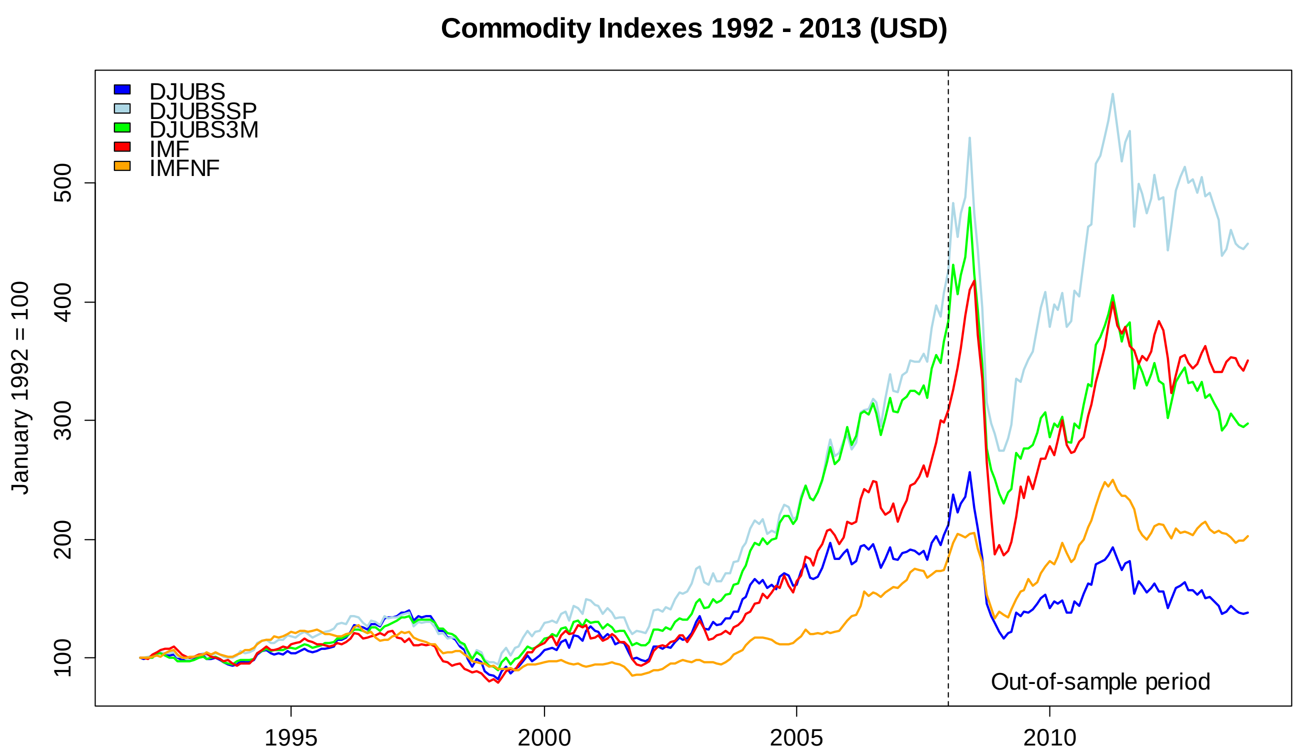
	DJUBS		DJUBSSP		DJUBS3M	
Last day	46.57		97.33		95.32	
Mean	45.52	(2.31 %)	94.92	(2.54 %)	93.03	(2.31 %)
Median	45.31	(2.79 %)	94.72	(2.75 %)	92.67	(2.86 %)

The least volatile were the median series. Their standard deviations were nearly 3 % less than in the last value of the month series. Also the volatility of the median series was nearly half a percentage less than for the mean series. Thus the median series were chosen for the models. While choosing the median series over the last value of the month removes some of the noise, typical of all financial markets, obviously the downside is discarding the most recent observations. This trade-off between reducing noise and foregoing the most recent value cannot be avoided. In this case**,** the most recent month was never found significant in explaining inflation, so sacrificing the most recent day of observation seems justifiable with hindsight.

4.2 Testing for Stationarity

Before any modeling can be done, the first step is to determine the order of integration and to check the stationarity of the time series. Among the reasons for this is avoiding incorrectly specified models such as spurious regressions, and determining whether the data should be transformed by differencing. There are two levels of stationarity, of which the more common is weak stationarity which demands that the mean and variance of the series are finite and constant and that all the autocovariances are independent of time. The more stringent definition is strict stationarity which requires that all the moments of the series, and not only the first and second (mean and variance), are unaffected by time (Verbeek, 2004). Causes of nonstationarity include unit-roots and structural breaks such as regime changes. The investigation here is concerned only with weak stationarity, and further mentions of stationarity always refer to weak stationarity.

Figure 2



Most statistical models cannot be correctly estimated from nonstationary data. Thus before any modeling it is required that the order of integration, i.e. the existence of one or more unit roots, is investigated, and if needed the data is transformed in the way needed to make it stationary. The most widely used test for stationarity is the Dickey-Fuller test and its augmented forms. Other common tests of stationarity include the Phillips-Perron (PP) test which is a derivative of the augmented Dickey-

Fuller (ADF) test, and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test where the null hypothesis is stationarity, while in the Dickey-Fuller, and PP tests the null is nonstationarity.

To determine the order of integration or stationarity of the inflation and commodities augmented Dickey-Fuller tests with 1, 2, 3, 6 and 12 lags, a short and long version of the Phillips-Perron test, and short and long versions of the KPSS test were performed. The test results are listed in Table 4. The expected result is for all the series to be integrated of the first order: I(1); in other words to have one unit root. For the commodities the test results are conclusive and consistent; they all point towards the commodities being nonstationary. In the case of inflation the ADF and PP tests reject stationarity, but the KPSS test rejects nonstationarity. To resolve the matter inflation will be tested more extensively.

Table 4

Stationarity Tests on Levels

	ADF 1	ADF 2	ADF 3	ADF 6	ADF 12	PP S	PP L	KPSS S	KPSS L
Inflation	0.55	0.53	0.33	0.46	0.41	0.49	0.43	0.1	0.1
DJUBS	0.97	0.99	0.99	0.97	0.95	0.98	0.97	0.01	0.01
DJUBSSP	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.01	0.01
DJUBS3M	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.01	0.01
IMF	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.01	0.01
IMFNF	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.01	0.04

For ADF and PP tests 0.99 implies p-value greater than 0.99
For KPSS test 0.1 implies p-value greater than 0.1 and 0.01 less than 0.01
S and L refer to short and long specification in PP and KPSS tests

Since performing the ADF, PP and KPSS tests on the inflation from 1992-2007 did not find a conclusive agreement on stationarity the testing is continued by increasing the sample to cover the entire available history of inflation from Statistics Finland, which is from January 1980 to March 2014, 411 months in total. KPSS tests are repeated on the history of inflation from 1980 to 2014, and now the p-values are less than 0.01 on both the long and short specifications of the test. Thus the stationarity of inflation can be conclusively rejected.

Since the R-squared of a unit-root nonstationary series converges to 1 as the sample grows to infinity (Tsay, 2005), we can use this knowledge to erase all doubt that inflation has a unit root. We estimate AR(1) models on the inflation series, and compare the estimation results to AR(1) models of nonstationary random walks. The random walks are simulated from cumulative sums of normally distributed random values.[8] The results are listed in Table 5 and in both cases; the entire inflation history of 411 months and the in-sample estimation 192 month period, inflation appears similarly unit-root nonstationary as the simulated random walk series. The R-squared values of the inflation AR(1) models clearly converge to 1 proving nonstationarity.

Table 5

AR(1) Model Without Constant					
	Estimate	**Std. Error**	**T value**	**R-squared**	**N**
Inflation 1980-2014:3	0.993	0.004	251.1	0.994	410
Random Walk N=411	0.989	0.007	132.9	0.977	410
Inflation 1992-2007	0.985	0.012	82.3	0.973	191
Random Walk N=192	0.984	0.014	72.9	0.966	191

Having established that the commodities and inflation are nonstationary, the next step is to determine their order of integration by repeating the same tests on the first differences. For the commodities the series are log-differenced, and inflation is differenced directly. The test results are shown in Table 6. From the test results can be concluded with relative confidence that all of the series are integrated of the first order, I(1). For inflation and DJUBS all of the tests are consistent in their finding that the differenced series are stationary. The case for the remaining four commodity indexes is not in total agreement: the PP tests and the ADF tests with 6 lags or less indicate stationarity, while the KPSS tests and the ADF test with 12 lags indicate nonstationarity.

[8] R code for random walk series reproduction:
set.seed(15); randomWalk411 = cumsum(rnorm(411)); set.seed(13); randomWalk192 = cumsum(rnorm(192))

Table 6

Stationarity Tests on Differences

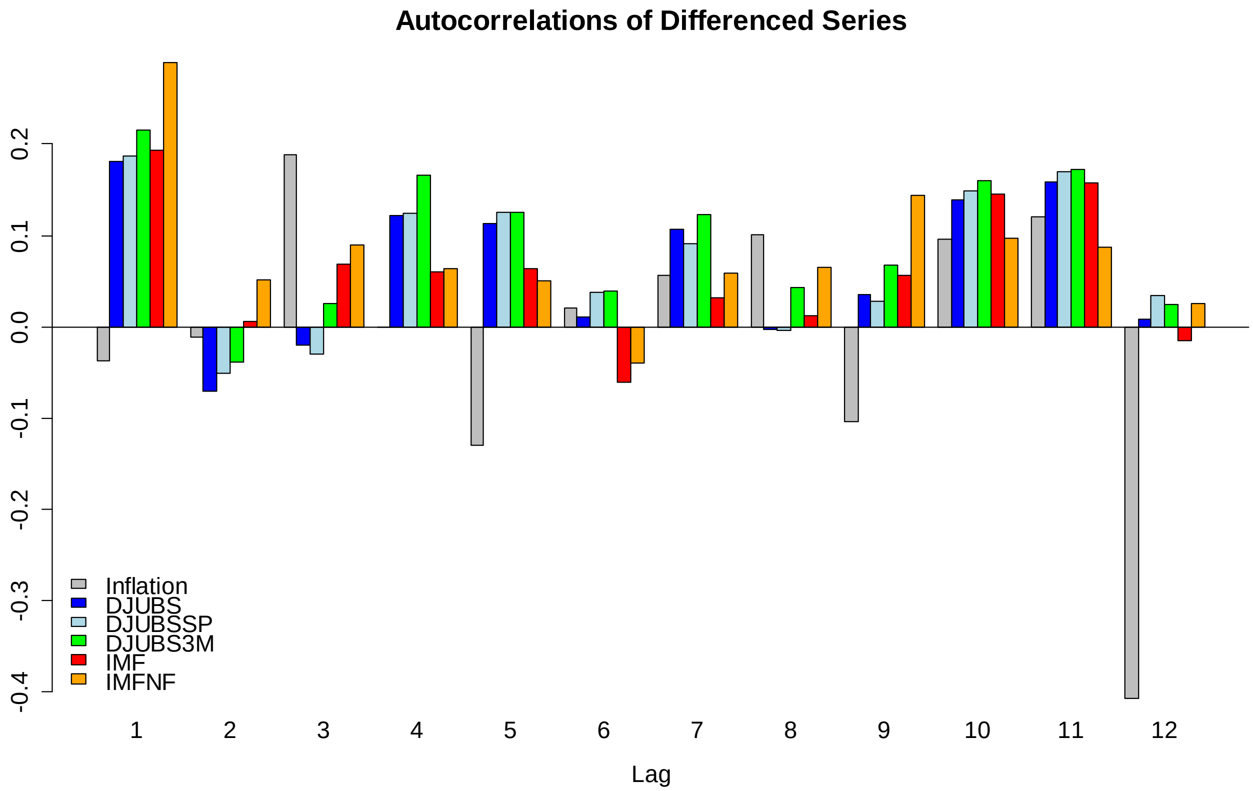
	ADF 1	ADF 2	ADF 3	ADF 6	ADF 12	PP S	PP L	KPSS S	KPSS L
ΔInflation	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.1	0.1
ΔLOG DJUBS	0.01	0.01	0.01	0.01	0.23	0.01	0.01	0.08	0.1
ΔLOG DJUBSSP	0.01	0.01	0.01	0.01	0.2	0.01	0.01	0.01	0.03
ΔLOG DJUBS3M	0.01	0.01	0.01	0.01	0.25	0.01	0.01	0.01	0.02
ΔLOG IMF	0.01	0.01	0.01	0.01	0.08	0.01	0.01	0.02	0.03
ΔLOG IMFNF	0.01	0.01	0.01	0.01	0.16	0.01	0.01	0.03	0.06

For ADF and PP tests 0.99 implies p-value greater than 0.99
For KPSS test 0.1 implies p-value greater than 0.1 and 0.01 less than 0.01
S and L refer to short and long specification in PP and KPSS tests

The identification of all the series as I(1) is chosen, despite slight disagreement remaining in the test results. First of all it would be difficult to accept that the rates of changes were not stationary for any of the series. A nonstationary rate of change would imply wildly erratic behavior both for inflation and the commodity prices. Secondly all the commodities must have the same order of integration; both their correlations in levels and rates of change (log differences) are so high, that the same order of integration really has to hold. The tests were consistent that DJUBS is I(1), but the same index consisting of spot-like prices, DJUBSSP, had mixed test results. However since the only significant difference between the DJUBS and DJUBSSP is that return effects from future rollovers have been removed from the latter, it is difficult to see how the spot series could have a higher order of integration.

To provide further illumination the first to twelfth autocorrelations of the differenced series are plotted in Figure 3. From it can be seen that the differenced series of inflation still has significant autocorrelation at the 12th lag, but the autocorrelations of the commodities exhibit no signs supporting nonstationarity.

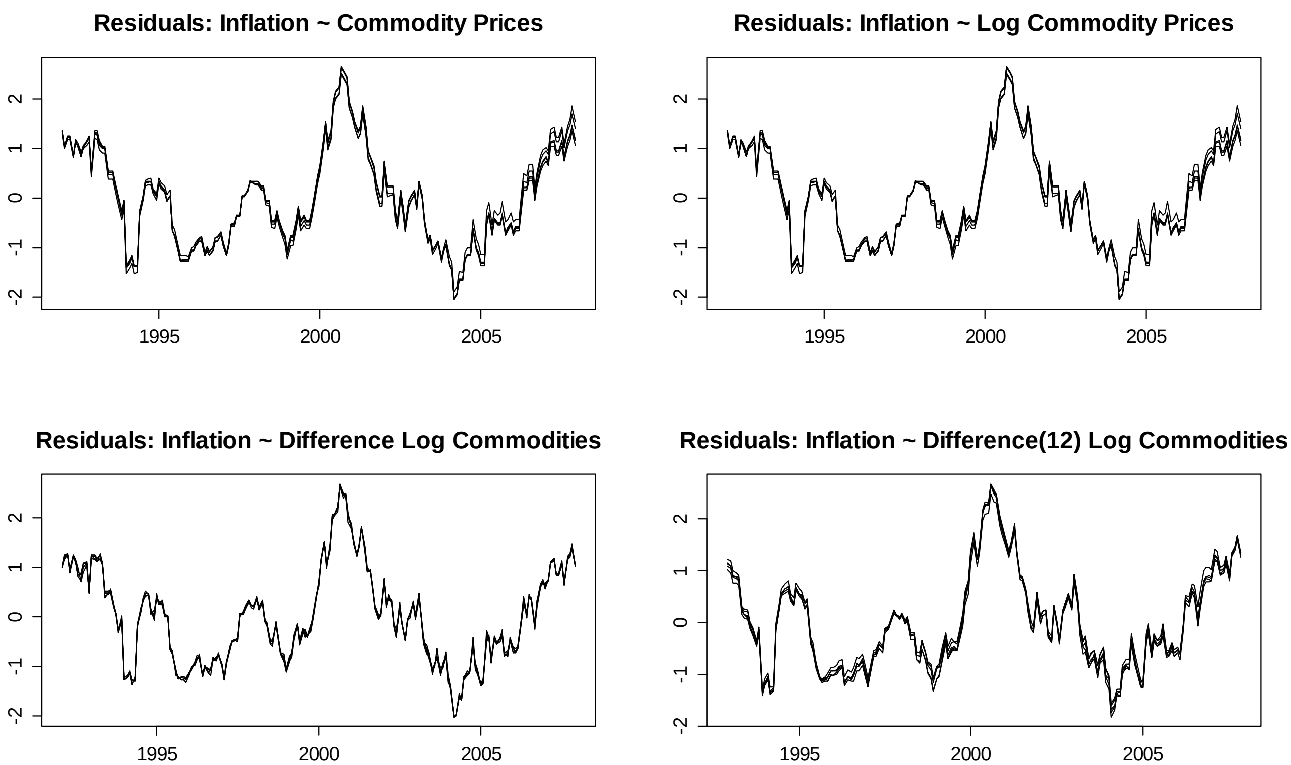
Figure 3



With all the series designated as I(1) what remains is testing for cointegration, or expressed differently whether the commodities and inflation share a common unit root. The rate of inflation and the level of commodity prices have in some studies been found to have a cointegrating relationship (example given Blomberg and Harris, 1995). Since all the commodity indexes and inflation were found I(1), we test for cointegration between inflation and all of the five commodity indexes with the Engle-Granger 2-step method. The tests are simply on the residuals produced by OLS regressions of inflation on the different commodities. To be thorough four different cointegrating relations are tested: the level of inflation with the level of commodities, the level of commodities in logs, the rate of change of commodities and finally the annual rate of change of commodities.

Each of the four tests produces five sets of residuals, which are depicted in Figure 4. A look at the residuals confirms that all of them are nonstationary. The same conclusion is further confirmed by ADF tests. In addition a cointegrating relationship would result in high R-squared figures for the regressions, but in fact the statistical significance of the commodity coefficients was rather low (with R-squared's no greater than 0.05). It seems quite clear that the unit-root of inflation is unrelated to the commodities, and that a linear combination of inflation and commodities producing stationary residuals does not exist. A cointegrating relationship between the levels of inflation and commodities can thus be conclusively ruled out.

Figure 4



5 Models

Forecasts are made at four horizons: 1 month, 3 months, 6 months and 1 year forward. The models specified are ADL models with the five different indexes, an AR model, a MA model and a model of the type specified by Atkeson and Ohanian. The ADL and AR models are fitted by ordinary least squares which is a consistent estimator for ADL models (Verbeek, 2004). The MA model is fitted by conditional least squares, and the AO model is just a simple moving average so it does not require ”fitting” as such.

The models are first fitted by minimizing the Akaike Information Criterion (AIC), after which the autocorrelation of the residuals is looked at and additional lags are added if warranted. Besides the AIC, there are other commonly used criterions used for fitting models. A similar alternative of the AIC is the Schwarz-Rissanen Information criterion or the Bayesian Information Criterion (BIC), which leads to similar but more parsimonious results. The AIC and BIC differ[9] in that the BIC penalizes more heavily for extra parameters.[10]

While an extensive agreement on which is better doesn't seem to have been reached, there is some evidence that with inflation lag selection with AIC outperforms BIC in out-of-sample forecasting performance due to inflation having a large moving average component (Edelstein, 2007). Generally in inflation literature use of AIC seems more common than BIC. The advantage that BIC has is that its asymptotic probability for overestimating the true size of the model is zero, while for AIC this probability is positive (Webb, 1988). With the trade off of possible model overestimation vs. forecasting performance, the latter was deemed more important, resulting in the AIC being selected. An additional advantage for the AIC is that in some cases of the ADL models the significance of commodity coefficients was rather weak, and possibly the BIC would have suggested removing them altogether, obviously complicating the model specification procedure.

For the models fitted by OLS another possibility would be fitting them according to the Adjusted R-squared, but this is an uncommon approach in a time series setting, and would result very likely to an overestimated model. Likewise it is possible if not practical to fit the univariate models directly from an examination of the autocorrelations and partial autocorrelations, but this offers no obvious advantage over using the AIC. The statistical significance of the coefficients also needs some attention paid, but they primarily served as a tool for pruning the model specification. Since the aim is to fit a model which explains the change of inflation as well as possible the individual p-values of the coefficients are not so important. If the model specified by AIC had some coefficients with e.g. a p-value of 0.12 it would not be seen as an issue even though the normal 0.05 significance level would not be achieved.

The AR, ADL and MA models were fitted with a constant even though the estimated values for the constants typically were very close to zero. A practical consideration for the inclusion of a constant is

---

[9] Formulas from Verbeek (2004): $AIC = \log \frac{1}{N}\sum_{i=1}^{N} e_i^2 + \frac{2K}{N}$, $BIC = \log \frac{1}{N}\sum_{i=1}^{N} e_i^2 + \frac{K}{N}\log N$

[10] Since 2 will always be greater than the logarithm of the number of observations whenever there are more than at least 8 of them. 2 < log N when N is greater than exp(2).

that its inclusion allows the calculation of long run means from the models. Without a constant the long run mean is forced to zero, which is equivalent to the level of inflation remaining the same, which is an unnecessary constraint.

The AR and ADL models were fitted directly, meaning separate models were fitted for all the forecast horizons, with all the previous values known at the time. Thus the 1-month forecast was modeled with lagged values from the previous month backwards, the three month forecast with values from the 3 previous months backwards and analogously for the six months and 1 year case. In contrast the MA model is fitted only once (for the next month), and all the longer forecasts are estimated by iterating this model recursively.

Marcellino et al. (2005) have found that recursive estimation can outperform direct estimation in forecasting performance. Using a VAR model it would be possible to estimate the commodity models also recursively, however unlike the change in inflation which has high autocorrelations as well as seasonality and is thus somewhat forecastable, the returns of commodity indexes behave more like random walks and are not forecastable in any useful sense. If they were since the DJUBS is the basis of numerous investment vehicles such as exchange traded funds (ETF's), the situation would violate the efficient market hypothesis, and be rather curious. Also if it was plausible that the Granger causality between Finnish inflation and commodity prices was bi-directional, the case for VAR modeling would be more justifiable. But the likelihood of Finnish inflation influencing world commodity prices is slight, since even US inflation has been found to not have causality on commodities (Acharya et al., 2010). A possibility would be following the examples of Furlong and Ingenito (1996) and Cecchetti et al. (2000) and using the true ex-post values of the commodity indexes as "forecasts", but this assumption is uncomfortable due to it being in near total opposition of reality.

Thus due to the one-sided directionality and the difficulty of forecasting commodity prices, which would not allow for any useful recursive estimations the VAR framework is not chosen, and the single equation ADL models are used.

5.1 Model Equations

The MA model specified: $\Delta\pi_t = \mu + \varepsilon_t + \alpha_1\varepsilon_{t-8} + \alpha_2\varepsilon_{t-10} + \alpha_3\varepsilon_{t-12}$ was found to have moving average lags at 8, 10 and 12. Since the MA model will be iterated for the 3 and 6 months and 1 year forecasts it is convenient that the lags specified; 8, 10 and 12, are large. Since all the lags are larger than 3 and 6, the three and six month iterations will be within the models scope and be estimated from actual errors (instead of estimated zero errors). Thus only the 1 year horizon forecasts will include iterated zero errors. Any forecast for a horizon beyond 1 year is simply the series mean μ .

Besides the MA model an ARMA model was also estimated, but the AIC lag selection led to a model which was nearly identical to the AR model, and for which the improvement of AIC was entirely marginal at only 0.2 % less than the AR models. The only difference between the AIC specified AR and ARMA models was that the formers AR lag at 8 being switched to a MA lag. Since the improvement in fit was so slight, and its similarity to the AR model so high, we felt that the ARMA model was not interesting enough for further discussion. In the out-of-sample forecasts the ARMA model was worse than both the AR and MA models.

The AO model: $\pi_{t+j} = \frac{1}{12}\sum_{i=0}^{11}\pi_{t-i} + \varepsilon_{t+j}$ is simply an arithmetic average of the 12 most recent months, and where j stands for the forecast horizon in months. For the forecasts on longer horizons the moving average window of the twelve observations is moved backwards the corresponding amount, e.g. for the 3-month forecast the average is counted from the 3-15 most recent months. Following the original model of Atkeson and Ohanian (2001), the inflation is modeled at levels. The original AO model models the level of quarterly inflation, and here the 1 year forecast model is analogous to the original AO model (with the exception of averaging the last 12 months instead of the last 4 quarters). The models for the 1, 3 and 6 month horizons are adapted versions of the 1 year model, in that 12 month window for the averaged inflations is moved forward the corresponding amount. As noted by Stock and Watson (2006) there is some ambiguity on the correct formulation of the AO model for forecast horizons other than a year, so others interpretations of the AO model are also possible.

The equations for the AR models are listed in Table 7, and the equations for the ADL models are listed in Table 8. Table 8 lists the equations for the DJUBS index, but the models for the other commodity indexes are equivalent. The fitted AR and ADL model lags are listed in Table 9. The AR lags specified for the 1 and 3 month models were the same: 3, 5, 8 and 12. Somewhat surprisingly the most recent 1 and 2 month AR lags were not explanatory enough to merit inclusion. The AR lags chosen for the 6 month forecast were 8, 11, and 12, and for the 12 month forecast only the 12th lag was specified.

Worthy of noting is that the 6 month model includes the 11th lag, which was not included in the 1 and 3 month models. Also the AR lags are the same for all the AR and ADL models, with the sole exception being the IMF 6 month model were the 11th AR lag was omitted.

Table 7

AR Models	
1 month:	$\Delta\pi_t = c + \theta_1\Delta\pi_{t-3} + \theta_2\Delta\pi_{t-5} + \theta_3\Delta\pi_{t-8} + \theta_4\Delta\pi_{t-12} + \varepsilon_t$
3 months:	$\Delta\pi_t = c + \theta_1\Delta\pi_{t-3} + \theta_2\Delta\pi_{t-5} + \theta_3\Delta\pi_{t-8} + \theta_4\Delta\pi_{t-12} + \varepsilon_t$
6 months:	$\Delta\pi_t = c + \theta_1\Delta\pi_{t-8} + \theta_2\Delta\pi_{t-11} + \theta_3\Delta\pi_{t-12} + \varepsilon_t$
1 year:	$\Delta\pi_t = c + \theta\Delta\pi_{t-12} + \varepsilon_t$

Table 8

ADL Models (DJUBS as an example)	
1 month:	$\Delta\pi_t = c + \theta_1\Delta\pi_{t-3} + \theta_2\Delta\pi_{t-5} + \theta_3\Delta\pi_{t-8} + \theta_4\Delta\pi_{t-12}$ $+ \phi_1\Delta\log DJUBS_{t-2} + \phi_2\Delta\log DJUBS_{t-3} + \phi_3\Delta\log DJUBS_{t-5} + \phi_4\Delta\log DJUBS_{t-7} + \varepsilon_t$
3 months:	$\Delta\pi_t = c + \theta_1\Delta\pi_{t-3} + \theta_2\Delta\pi_{t-5} + \theta_3\Delta\pi_{t-8} + \theta_4\Delta\pi_{t-12}$ $+ \phi_1\Delta\log DJUBS_{t-3} + \phi_2\Delta\log DJUBS_{t-7} + \phi_3\Delta\log DJUBS_{t-13} + \varepsilon_t$
6 months:	$\Delta\pi_t = c + \theta_1\Delta\pi_{t-8} + \theta_2\Delta\pi_{t-11} + \theta_3\Delta\pi_{t-12}$ $+ \phi_1\Delta\log DJUBS_{t-7} + \phi_2\Delta\log DJUBS_{t-13} + \varepsilon_t$
1 year:	$\Delta\pi_t = c + \theta\Delta\pi_{t-12} + \phi_1\Delta\log DJUBS_{t-13} + \phi_2\Delta\log DJUBS_{t-21} + \varepsilon_t$

The commodity lags chosen for the ADL models varied considerably. The lags at 3, 7 and 13 were common and present in most of the models if applicable. The rest of the lags varied quite a bit. Again curiously the most recent 1 month lag was not included in any of the 1 month models. While it could be easily believed otherwise, the commodities have no immediate effect on Finnish inflation. It appears that not even energy prices, the increases of which are commonly believed to be passed on swiftly to the prices of fuels, do not have an immediate effect on Finnish inflation. The nearest commodity lags included are the 2nd in the DJUBS and DJUBSSP models. Curiously while the 1st lag is never included the 13th one is often selected. Since the models are fitted in such a way that only the most recent 12 lags are considered, it is not investigated whether or not the 13th lag would be selected into the 1-

month forecast as well.

Table 9

AR and ADL Model Lags

	Lags:	Inflation	Commodity
1 month	AR	3 5 8 12	-
	DJUBS	3 5 8 12	2 3 5 7
	DJUBSSP	3 5 8 12	2 3 5 7
	DJUBS3M	3 5 8 12	3 5 7
	IMF	3 5 8 12	3 5 7 11
	IMFNF	3 5 8 12	3 5 7
3 months	AR	3 5 8 12	-
	DJUBS	3 5 8 12	3 7 13
	DJUBSSP	3 5 8 12	3 5 7
	DJUBS3M	3 5 8 12	3 4 9 13
	IMF	3 5 8 12	3 7 11 13
	IMFNF	3 5 8 12	3 5 7
6 months	AR	8 11 12	-
	DJUBS	8 11 12	7 13
	DJUBSSP	8 11 12	13
	DJUBS3M	8 11 12	7 13
	IMF	8 12	11 13
	IMFNF	8 11 12	7
1 year	AR	12	-
	DJUBS	12	13 21
	DJUBSSP	12	13
	DJUBS3M	12	13
	IMF	12	13 18 21
	IMFNF	12	21

5.2 Summary Statistics

The model residuals were tested for autocorrelation, heteroscedasticity and normality. The results are listed in Table 10. Residual autocorrelation was tested with Ljung-Box tests run with 12 lags which did not find remaining autocorrelation in any of the models. Likewise the Breusch-Pagan tests did not locate heteroscedasticity, indicating that the linear formulation of the ADL models was correct. Ideally the residuals should follow a normal distribution with a constant variance. To evaluate normality the Shapiro-Wilk and Jarque-Bera tests were both done on the residuals. The results were mixed since the

Shapiro-Wilk tests did not reject normality in any of the cases, while the Jarque-Bera tests had small p-values which are significant at the traditional level of significance 0.05. The conclusion to be drawn is that while the mean and variance of the residuals satisfies normality as tested by Shapiro-Wilk test, the higher moments, skewness and kurtosis, examined by the Jarque-Bera test do not. As judged by the Jarque-Bera test the 3-month forecast models are the best behaved, and as the forecast horizon increases, the normality of the residuals deteriorates.

Table 10

Model Fit Diagnostics

	Adjusted R^2	Ljung-Box (12)	Breusch-Pagan	Shapiro-Wilk	Jarque-Bera	N
1 month						
AR	0.233	0.240	0.669	0.342	0.038	179
DJUBS	0.287	0.382	0.266	0.432	0.086	179
DJUBSSP	0.290	0.444	0.210	0.321	0.052	179
DJUBS3M	0.270	0.398	0.363	0.452	0.087	179
IMF	0.280	0.322	0.050	0.340	0.070	179
IMFNF	0.254	0.245	0.265	0.284	0.044	179
3 months						
AR	0.233	0.236	0.681	0.374	0.047	177
DJUBS	0.286	0.273	0.596	0.424	0.074	177
DJUBSSP	0.288	0.459	0.225	0.482	0.117	177
DJUBS3M	0.259	0.195	0.717	0.400	0.191	177
IMF	0.288	0.275	0.091	0.097	0.014	177
IMFNF	0.253	0.241	0.292	0.355	0.060	177
6 months						
AR	0.203	0.336	0.705	0.273	0.024	174
DJUBS	0.223	0.264	0.878	0.103	0.002	174
DJUBSSP	0.207	0.273	0.883	0.164	0.007	174
DJUBS3M	0.224	0.246	0.872	0.109	0.003	174
IMF	0.238	0.277	0.182	0.046	0.000	174
IMFNF	0.205	0.313	0.193	0.406	0.041	174
1 year						
AR	0.178	0.452	0.267	0.175	0.035	168
DJUBS	0.201	0.538	0.728	0.281	0.042	168
DJUBSSP	0.189	0.453	0.612	0.147	0.009	168
DJUBS3M	0.195	0.411	0.492	0.125	0.009	168
IMF	0.208	0.427	0.914	0.261	0.015	168
IMFNF	0.199	0.516	0.936	0.637	0.269	168

While there is no easily available remedy for the non-normal skewness and/or kurtosis of the residuals, it is important to acknowledge this finding to limit the confidence in forecasts drawn form the models.

The MA models residuals were also tested. The Ljung-Box p-value was 0.686, but both the Shapiro-Wilk (p-value 0.001) and Jarque-Bera (p-value < 0.0001) tests rejected normality for the residuals with highly significant p-values.

To check for unexplained seasonality or shifting dynamics the residuals were examined further. Since there are so many models only the residuals of only the best 1-month forecast model, as judged by Adjusted R-squared (which is DJUBBSP) and the MA model, are checked. In addition to the residuals also their squares are examined. The squared residuals are centralized around zero by subtracting their mean from them. For both sets of residuals, from the DJUBSSP and MA models, four OLS models are then fitted. Models with dummy variables either for months or years are fitted on both the residuals and their squares. All of the models are fitted without a constant since the mean of both kinds of residual series is zero by definition.

An analysis of variance (ANOVA) was then done for all the eight models. In none of the cases was the p-value for the F-statistic significant. However while the ANOVA's did not find evidence for the residuals being different by month or year, some of the coefficients in the OLS models were significant or nearly so. In the models with the dummies for years, the squared residuals of the year 1993 were larger than average with a p-value smaller than 0.01 for both the MA and DJUBSSP models. This implies that the models perform worse for the year 1993 than in the other years from 1992 to 2008 and most likely some exogenous shock was behind this.

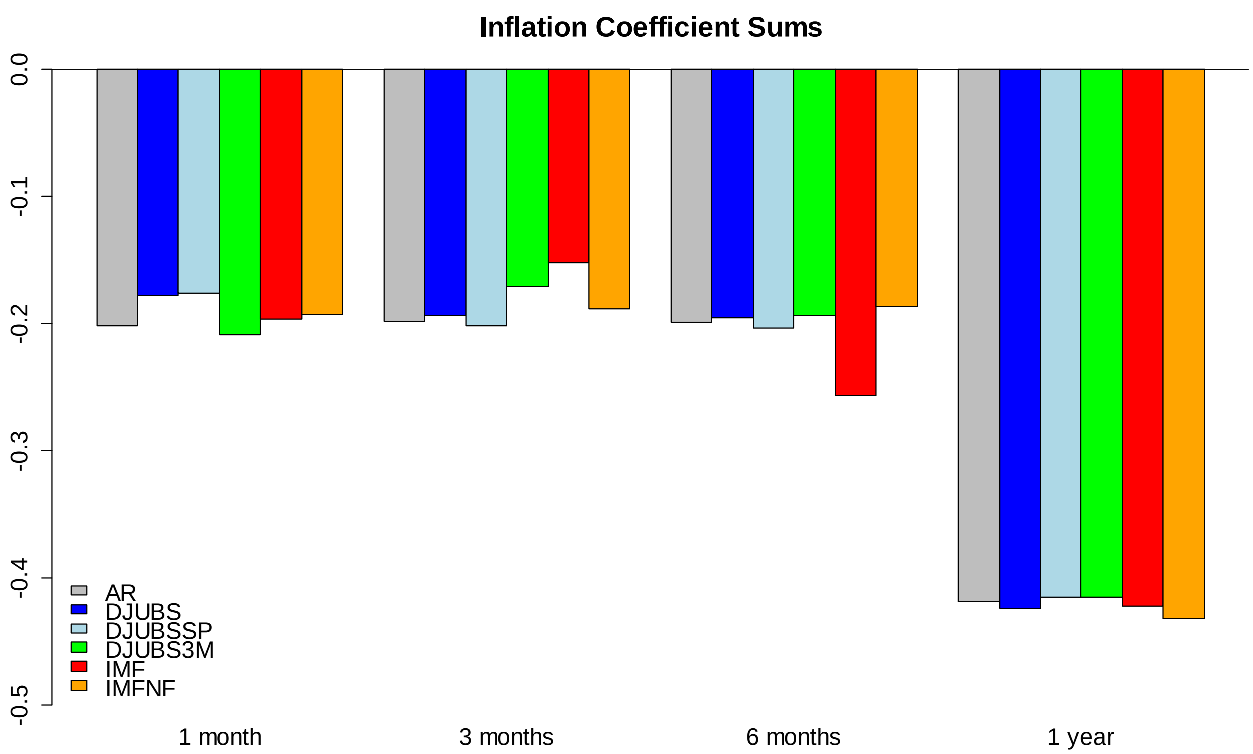
The month of January exhibits some deviating behavior as well. For the MA model January's coefficients in both the residual and squared residual models are significant at a p-value smaller than 0.01. Fortunately the DJUBSSP performs better with the p-values for January being 0.12 in the residual model and 0.09 in the squared residual model. January is the unique special case, since in none of the models were the coefficients for the other months even close to significant. Several plausible explanations could explain the different behavior in January, perhaps Christmas or the large scale post-Christmas sales in January. While including a dummy variable for January in the inflation models could well improve their fit, this approach will not be taken here since it hard to justify why the abnormal effect from January should be constant. Also the January effect was only really a problem for the MA model, and not significantly so for the DJUBSSP model.

5.3 A Look at the Coefficients

The sum of autoregressive coefficients is a common measurement for inflation persistence (Cecchetti and Moessner, 2008). With the ADL models here both the inflation persistence and the influence from the changes in the past commodity prices are explored. The primary point of interest here are the long-run multipliers of the commodities, but to start of we check whether the sums of the coefficients for the lags of inflation are of similar magnitudes.

The sums of the AR and ADL models inflation lag (AR) coefficients are depicted in Figure 5. What is noticeable is that the sums of the inflation coefficients at the different forecast horizons are similar in scale, whether or not the AR model is augmented with commodities. From this can be cautiously inferred that the information the commodities bring is not already contained in past inflation. If the information from past inflation and the commodities was mostly the same, the models would have experienced multicollinearity likely leading to more variance in the AR coefficients.

Figure 5



The inflation coefficient sums at one, three and six months are of similar size, while considerably less at 12 months. This difference is due to the one year models only lag of inflation being the 12th lag. The 12th lag is more a seasonal element than a purer expression of causality from past inflation. While there can be true seasonality in prices, at least partially this seasonality arises from the definition of inflation since inflation is the change in consumer prices from a year ago. Thus controlling for seasonality by excluding the 12th lag from the sum, the sums at 1, 3, and 6 months should be around 0.2. For the AR model these sums were 0.194, 0.200 and 0.218 for the 1, 3 and 6 month models. Notably the sign of the sum of coefficients changes, and is positive, when the 12th lag is excluded. The positive coefficient is as expected since the changes in inflation should exhibit positive autocorrelation, since it can have temporary (but long) trends of either decreasing or increasing.

We now proceed to the long-run multipliers of the commodities which hopefully provide some insight on inflation dynamics. The long-run multiplier is simply defined as the sum of the commodity coefficients divided by (1 – the sum of the inflation coefficients). To see why this holds the ADL equation $\Delta\pi_t = c + \sum \theta_i \Delta\pi_{t-i} + \sum \phi_j \Delta \log X_{t-j} + \varepsilon_t$ is arranged into its long-run mean form:

$$\Delta\pi = \frac{c}{1-\sum\theta_i} + \frac{\sum\phi_j}{1-\sum\theta_i}\Delta\log X \ .$$

The derivation is detailed in Appendix 2. From the long run mean equation can be seen that the long run change in inflation consists of two parts, the first one is simply the long run mean of an AR model, and the second is the long-run multiplier of the commodities rate of change.

If modeled from the levels of inflation and commodity prices the interpretation of the multiplier would be straightforward, simply expressing the permanent relation between those levels. Since we are modeling rates of change instead of levels the interpretation is different. The long-run mean of inflation's rate of change has to be zero; otherwise the result would be an explosion in the level of inflation. Thus here, despite its name, the effects of the long-run multiplier are transitory and not permanent. Even though the effect here is only transitory, we choose to retain the "long-run multiplier" designation, since as is seen from Appendix 2 the derivation of the multiplier is dependant on adopting the long-run view.

Since here the multipliers are of transitory and not permanent effects, they can differ at different forecast horizons. Following the “Tortoise and Hare” theory of Blomberg and Harris (1995), we expect the multipliers to be positive at shorter horizons, and reverse to negative at some longer horizon. Thus both the signs as well as the magnitudes of the multipliers are interesting. Also the horizon at which the multipliers sign turns negative is interesting. We estimate models only at four horizons, but a completer picture of the multipliers at different horizons could be obtained by estimating more models (e.g. twelve for every horizon from 1 to 12 months).

Both cases, negative and positive multipliers have economically feasible explanations. The more straightforward case of positive multipliers implies that there is a direct relationship between the rates of change of commodity prices and inflation, such as a rise in the price of oil, leading to higher prices for products like gasoline, and more expensive manufacturing and transportation. The direct effect is the second of the theories outlined by Blomberg and Harris (1995). A notorious example of this direct effect was the stagflation era of the 1970s when in the industrial countries an oil supply shock resulted in an increase in the price of gasoline, and drove everything else more expensive as well.[11]

This effect could have either an instant or delayed effect or a first or second round effect as they are often called. For example adjustments to the price of gasoline should be rapid, while the effects from manufacturing becoming more expensive should be more delayed. Also the third theory by Blomberg and Harris (1995) of investors using commodities as an inflation hedge is based on the correlation between commodities and inflation being positive.

Negative multipliers should arise mainly from the latter phase of the Tortoise and Hare dynamic, where the initial over-adjustment reverses. Important to remember is that the relation between commodities and inflation can also be confounded by exchange rates and monetary policy (Blomberg and Harris, 1995). Thus negative multipliers could possible arise also from central bank policy, if the central bank overreacts to inflationary pressures from commodities, and ends up more than neutralizing them, reversing the tide of inflation. But a negative multiplier has to be transitory; consumer and commodity prices cannot be moving in opposite directions for any extended period.

The sums of the coefficient long-run multipliers are listed in Table 11 and plotted in Figure 6. The inferences that can be made are limited. Notably the IMFNF which does not contain energy prices

[11] http://www.investopedia.com/articles/economics/08/1970-stagflation.asp

varies less at different horizons. At horizons from one to six months, the multipliers are essentially the same with each approximately 0.6. If one wished to interpret this as a permanent long-run multiplier, it would mean that a permanent monthly growth rate of one percent of IMFNF results in monthly increases of inflation by 0.006, or 0.072 annually. Even as a permanent long-run relation the multiplier's scale is within the realm of what could be possible.

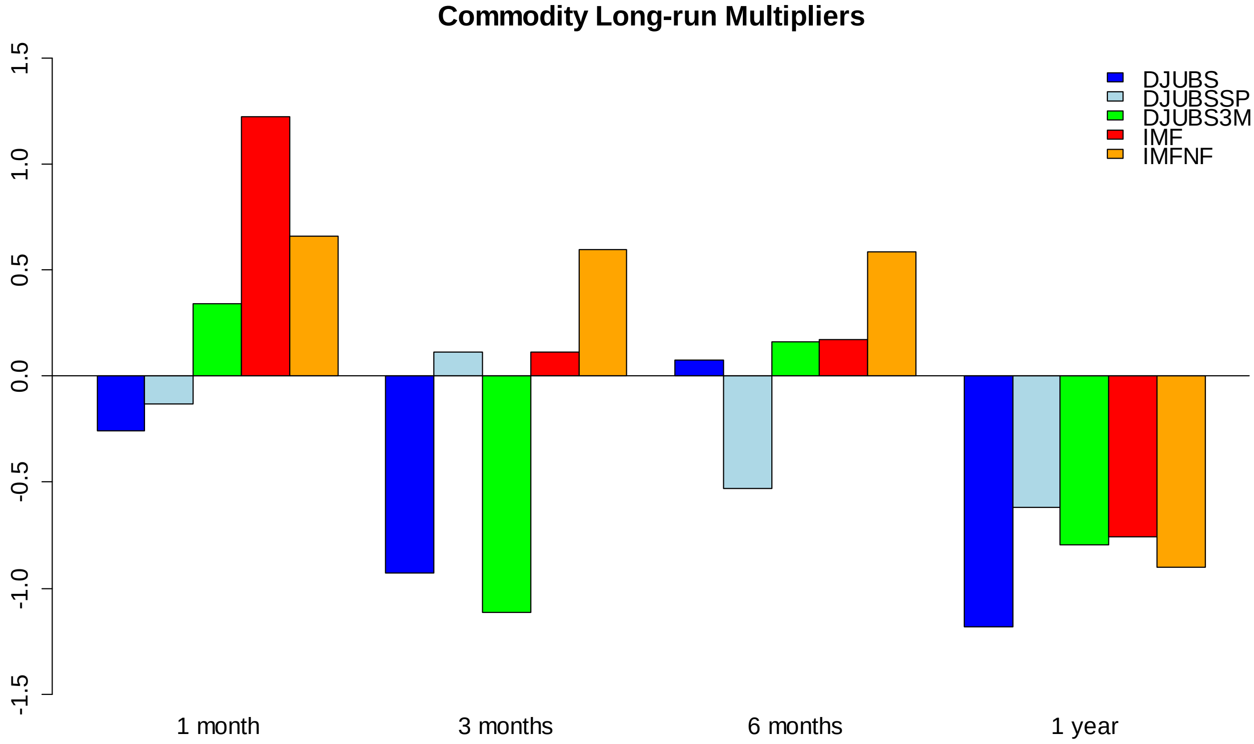
Table 11

Long-run Multipliers

	1 month	**3 months**	**6 months**	**1 year**
DJUBS	-0.26	-0.93	0.08	-1.18
DJUBSSP	-0.13	0.11	-0.53	-0.62
DJUBS3M	0.34	-1.11	0.16	-0.79
IMF	1.22	0.11	0.17	-0.76
IMFNF	0.66	0.60	0.59	-0.90

There is quite a bit of variation between the multipliers. However they all are consistent in that at the one year horizon the multipliers are all negative. Obviously there is no guarantee for all of them not being consistently wrong on this. Since in every case, at 1, 3, and 6 months models two fifths of the multipliers are negative and the other three positive, it does not seem likely that a reliable and well defined causality exists. Evaluated from the vantage point of the Tortoise and Hare dynamic, the Dow Jones models do not behave as expected, while the IMF and the IMFNF have the correct signs to fit the narrative. Obviously the credibility of the multiplier and the forecasting accuracy are related, so the credible interpretations from the multipliers are contingent on the ADL models outperforming their benchmarks decisively.

Figure 6



6 Results

The case for commodities serving a useful role in forecasting inflation does not look impressive. The root mean squared forecast errors (RMSFE) are listed in Table 12, and more often than not the ADL model with the commodities are outperformed by the benchmark AR model. None of the commodities is able to improve the one year forecast, but in some cases at the shorter horizons, most notably the energy-less IMFNF, they have a RMSFE less than the benchmarks. The significance of these improvements on a pure autoregressive model will be examined further on in this chapter.

Table 12

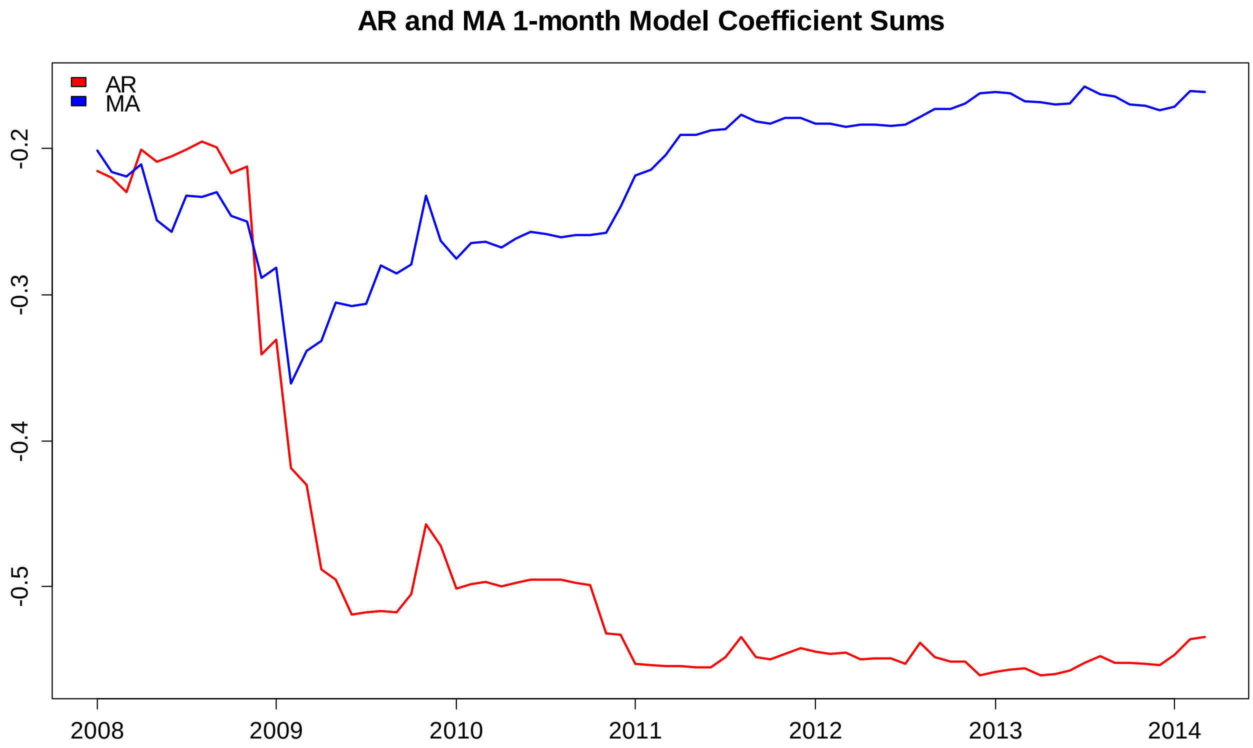
Root Mean Squared Forecast Errors				
	1 month	**3 months**	**6 months**	**1 year**
MA	0.335	0.338	0.342	0.337
AR	0.355	0.355	0.359	0.360
DJUBS	0.367	0.367	0.355	0.368
DJUBSSP	0.366	0.364	0.359	0.364
DJUBS3M	0.358	0.357	0.355	0.364
IMF	0.349	0.349	0.360	0.373
IMFNF	0.347	0.348	0.355	0.369
AO	1.367	1.682	2.059	2.475

Something else notable is that the performance of the AO model is nothing short of disastrous. During the period from 2008 to 2013 the standard deviation for the monthly change in inflation was 0.4, but even at the nearest forecast horizon of one month the RMSFE of the AO model is nearly three times as much. The AO models forecast error increases further as the forecast horizon becomes, reaching an incredible RMSFE of 2.5 % for the one year forecasts. There is a quite simple explanation for the models performance. Unlike the rest of the models, the AO model is on the level, and not the rate of change, of inflation. Thus its results with the other models are not strictly comparable. The AO model only functions well in periods where the inflation fluctuates around some stable mean, and this has not been the case from 2008 onwards.

Another observation to be made is the MA model considerably outperforming the AR model. Their performance is directly comparable only in the one month forecast, since at longer horizons the MA model is iterated, while the AR models are estimated directly without iteration. This is not surprising, since typically in the recent inflation literature MA models have been found to provide better out-of-sample forecasts than AR models (Stock and Watson, 2006). Additionally iterated forecasts have been found to forecast better than directly estimated ones (Marcellino et al., 2005).

Interestingly the AIC for the AR model was smaller than for the MA model. It is not certain what conclusion should be drawn from this, perhaps while the AR models fit is better (and thus the AIC smaller), it is a result of overfitting and capturing noise as well, more so than in the MA model. Alternatively it could be that the inflation generating process experienced a structural change during the forecasting period, and the MA models were less adversely affected by the change. Figure 7 depicts the sums of the coefficients from the AR and MA models, and it shows that these sums changed considerably during 2008 for both models. The MA coefficient sum later rebounds to a similar level with its initial January 2008 value. However the AR coefficient sum drops permanently and beginning in 2011 stabilizes to a much lower value than where it began.

Figure 7



Returning back to this studies central question of comparing the ADL models to their AR model benchmarks, the RMSFE's of the models were presented in Table 12. To illuminate the forecasts utility Table 13 presents the percentage of months where the respective commodity-models outperformed the AR model. As can be seen from Table 13 none of the ADL models was consistently at every horizon more often than not closer to the true value than the AR model. In a result consistent with their RMSFE's the IMFNF models were closer to the true value 56 % of the time at the 1 and 3 month horizons, but at the 6 month horizon despite the RMSFE of IMFNF being smaller than the benchmark the AR model is more often closer to the true value. Notably while none of the commodity models had a smaller RMSFE than the benchmark at the one year horizon, two of the DJUBS models, DJUBSSP and DJUBS3M, performed well here, with the DJUBS3M being closer 59 % of the time. Of all the models the one with the highest proportion of beating its benchmark was the DJUBS3M at 1 year. Even then the result could well be a result of pure chance. Assuming that the "true" probability for beating the benchmark is 50 %, in an out-of-sample period of 6 years and 3 months there are 75 months so based on the binomial distribution the likelihood of getting 59 % correct is about 5 %[12], which does not rule out pure chance by any means.

Table 13

Better than AR Model				
	1 month	**3 months**	**6 months**	**1 year**
DJUBS	48%	55%	48%	48%
DJUBSSP	48%	49%	57%	56%
DJUBS3M	47%	57%	49%	59%
IMF	53%	51%	45%	40%
IMFNF	56%	56%	47%	40%

While it seems that commodities did at least in some cases improve on the univariate forecast, to temper even this modest excitement the scale of the improvement must be noted. As listed in Table 14 the biggest winner compared to AR model is the IMFNF at one and three months; with improvements of 0.007. This is equivalent to 0.7 basis points which is 0.007 %. Keeping in mind that the precision in which inflation is published is ten basis points (or 0.1 %) an improvement of less than a basis point is insignificant.

Table 14

Difference to AR Model RMSFE				
	1 month	**3 months**	**6 months**	**1 year**
DJUBS	-0.012	-0.012	0.004	-0.008
DJUBSSP	-0.012	-0.010	0.000	-0.003
DJUBS3M	-0.003	-0.002	0.004	-0.004
IMF	0.005	0.006	-0.002	-0.013
IMFNF	0.007	0.007	0.004	-0.009

While the improvements on the AR forecast are indeed moderate, their statistical significance is tested next. The test performed is a paired one sided t-test, in other words it is tested whether the difference between the squared errors of the AR model and the commodity models are significantly greater than

---

[12] Cumulative binomial distribution: $1-\sum_{i=0}^{44}\binom{75}{i}0.5^{i}(1-0.5)^{75-i}$

zero. The same tests two-sided version was performed by Eugeni and Kruger (1994), but here the one-sided specification is chosen. The one-sided test is more stringent, but also more fitting since the question is whether the commodities improve on the AR forecast and not whether the forecasts are the same. The results of the t-tests are listed in Table 15.

Table 15

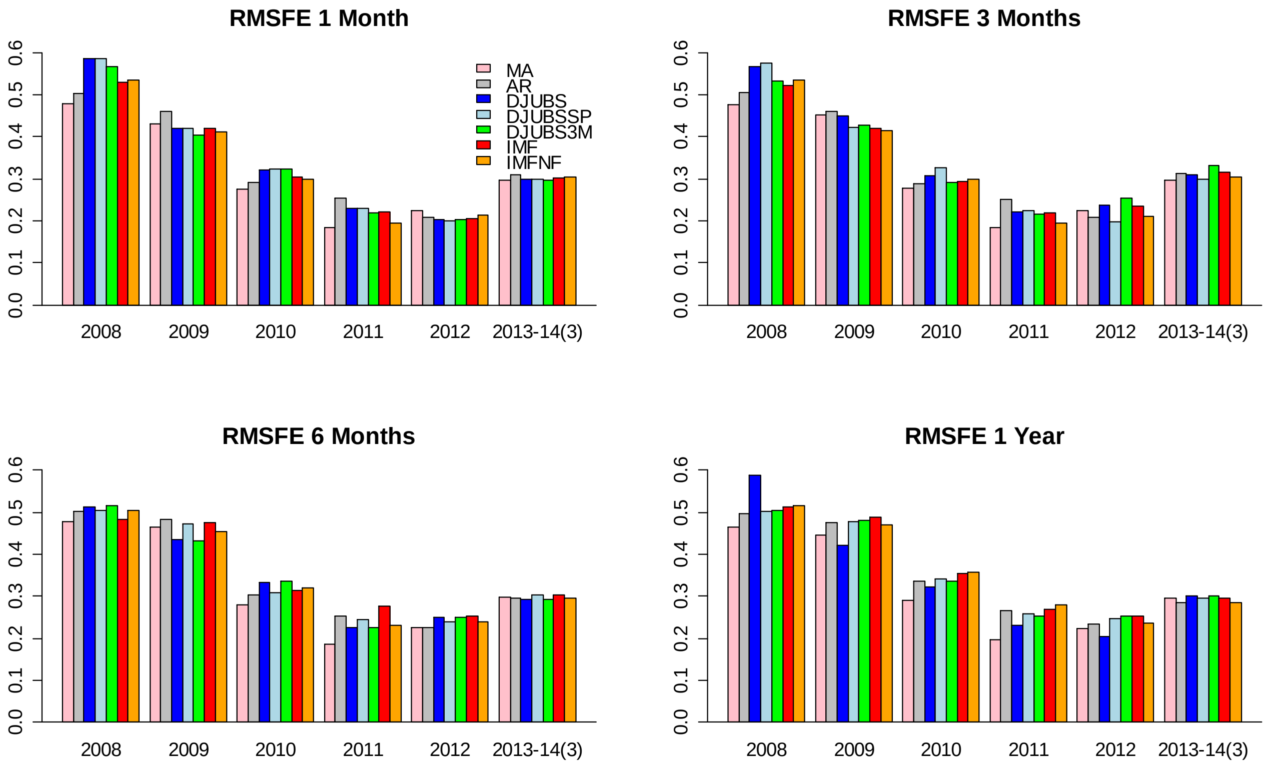
T-tests on Squared Forecast Error Differences

	1 month	**3 months**	**6 months**	**1 year**
DJUBS	0.796	0.838	0.318	0.864
DJUBSSP	0.772	0.737	0.486	0.731
DJUBS3M	0.580	0.567	0.348	0.730
IMF	0.311	0.291	0.562	0.950
IMFNF	0.261	0.288	0.252	0.900

For none of the models was the improvement on the AR model significant at the conventional 0.05 level. Nearest to this mark came the IMFNF model for the 6 month forecast with a p-value of 0.252. IMFNF was the only commodity to have several p-values smaller than 0.3, which it did at 1, 3 and 6 months. In the cases where the AR model had a lower RMSFE, the p-value is equivalently greater than 0.5.

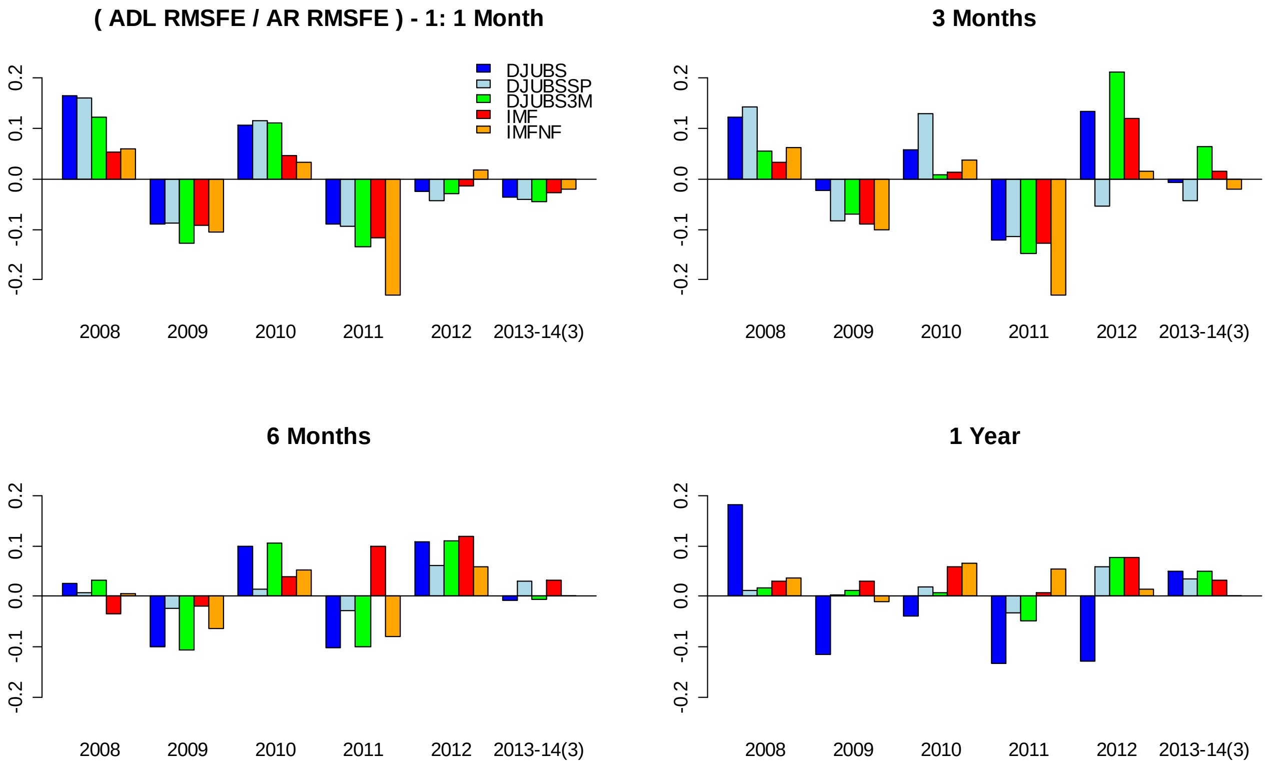
How consistent are the results with the passage of time is important to evaluate. Figure 8 depicts the RMSFE's of the models broken down to annual values for 2008 to 2013. Since there are only forecasts from up to March for 2014, these last three forecasts were appended to 2013. What stands out is that the performance of the models was much better 2010 onwards than in 2008 or 2009. For all the models the average error is below 0.4 % in the last 4 years, and significantly more before this. As can be seen from Figure 1 the level of inflation rocketed in 2008 and subsequently came down fast enough for deflation arising in 2009. Since then the situation has stabilized, and consequently our models performance has improved. Thus it seems that both the simple univariate AR and MA, as well as the ADL models are fair-weather models of sorts, they're reliable in "normal" settings, but less so in abnormal ones. This is acceptable for the univariate models, but for the commodity ADL models to really serve any useful purpose, they should improve on the univariate forecasts especially in changing or more volatile times.

Figure 8



The ratio between the commodity ADL models and the AR model's yearly RMSFE's are drawn in Figure 9. It does not show evidence of either the performance of the ADL models compared to the benchmark either improving or worsening during time.

Figure 9



Another further exploration on the robustness of the models' performances is looking at the correlation between the absolute values of the change in inflation and square roots of the forecast errors squares. These are listed in Table 16. The results indicate that the correlations are high, approximately 0.8 for every commodity model, which confirms the suspicions that the models performance negatively correlates with the level of volatility in inflation. This is not unsurprising or unacceptable; it is just that the opposite case of the forecast errors being uncorrelated with inflation volatility would be more advantageous for practical uses.

Table 16

Correlation: Abs(∆Inflation) & Abs(Forecast Error)				
	1 month	3 months	6 months	1 year
AR	0.816	0.814	0.845	0.844
DJUBS	0.774	0.784	0.793	0.813
DJUBSSP	0.772	0.772	0.846	0.823
DJUBS3M	0.767	0.796	0.781	0.814
IMF	0.782	0.760	0.800	0.802
IMFNF	0.785	0.786	0.800	0.832

6.1 Forecast R-squared

An alternative expression for the average forecast errors is to calculate an R-squared value from them. This is achieved by dividing the mean squared forecast error (MSFE) with the population variance of the inflation series.[13] The out-of-sample R-squared values are listed in Table 17 and shown in Figure 10. While they are simply scaled versions of the MSFE figures, they are perhaps more expressive. Arguably for example the difference between the 1 month R-squared between IMFNF and the AR, 0.207 vs. 0.173 is more dramatic than their 0.7 basis point difference in RMSFE was. The superiority of the MA model over the AR specification is also once again driven home; R-squared's of the AR models are around two thirds of the MA models.

Table 17

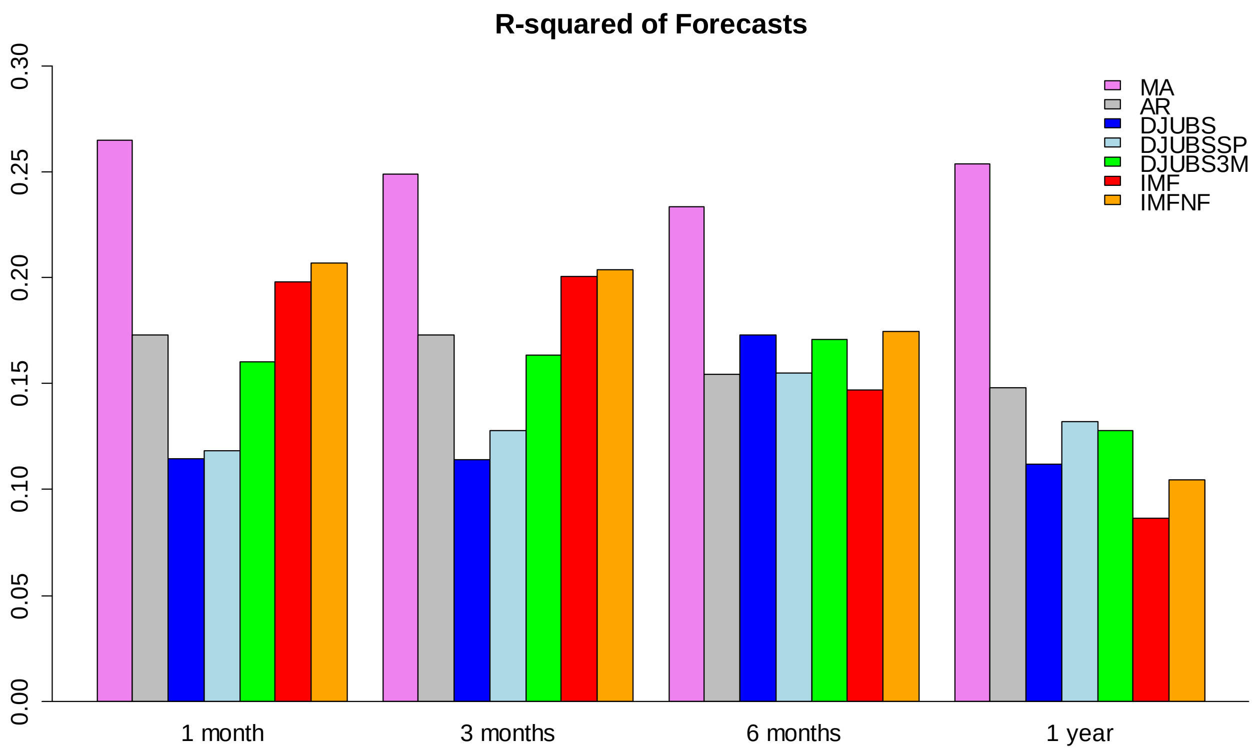
Forecast R-squared				
	1 month	3 months	6 months	1 year
MA	0.265	0.249	0.234	0.254
AR	0.173	0.173	0.154	0.148
DJUBS	0.115	0.114	0.173	0.112
DJUBSSP	0.118	0.128	0.155	0.132
DJUBS3M	0.161	0.163	0.171	0.128
IMF	0.198	0.201	0.147	0.086
IMFNF	0.207	0.204	0.174	0.105

---

13 $$R^2 = 1 - \frac{MSFE}{\frac{1}{T}\sum\left(\pi_t - \bar{\pi}\right)^2}$$

The motivation for using the population instead of sample variance formula is that MSFE since it is a mean is scaled by 1/T.

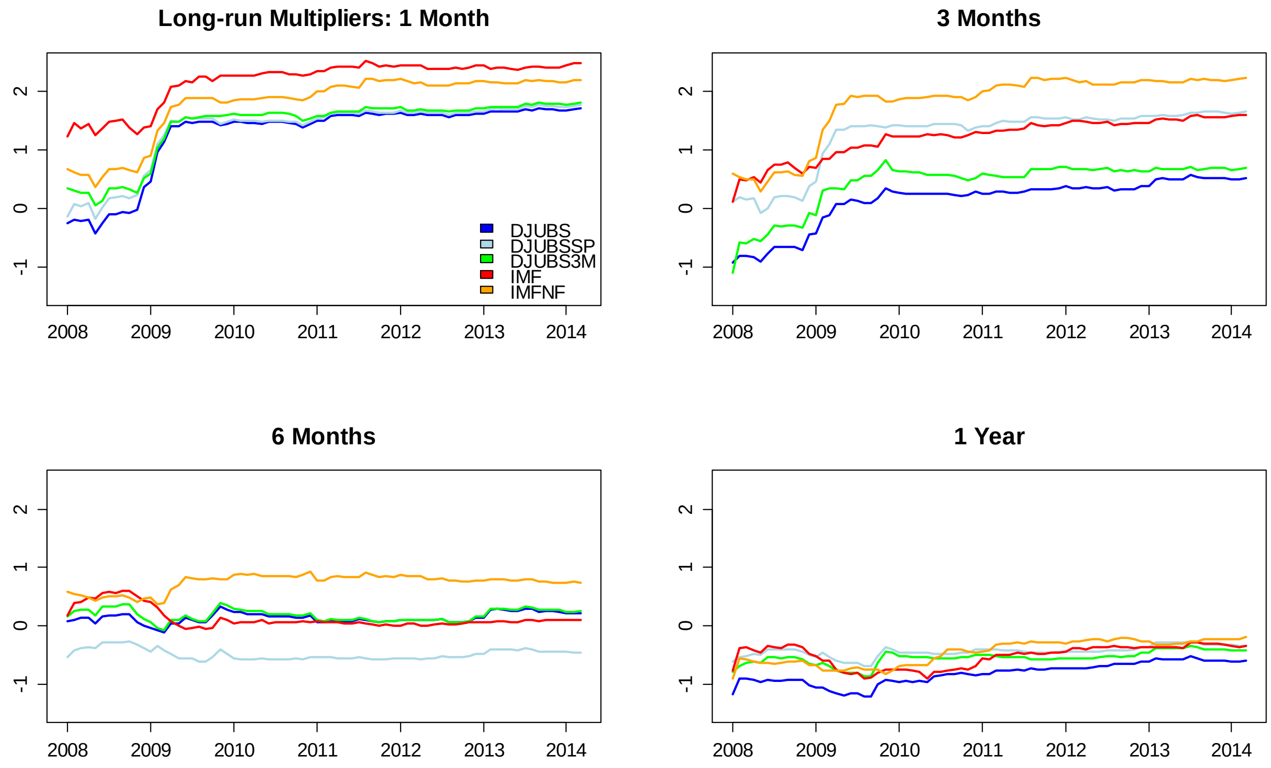
Figure 10



6.2 Coefficient Stability

To investigate how stable the commodity and inflation relation is, it serves to take a look at the evolution of the commodity coefficients during the out-of-sample period. The long-run multipliers as they evolve in the in the forecast period are depicted in Figure 11. They cannot be classified as stable, and they change a lot with time, at least during the year 2008. Most noticeably in the 1 month forecast plot can be seen how once the values from 2008 enter the estimation history the coefficient sum jumps to a higher level where it remains in a stable manner to the end of the forecast period in March 2014. To be noted is that the coefficients are from recursively estimated models, so the instability of the coefficients is even more striking. A rolling window for estimation should produce even less stable coefficients.

Figure 11



It is open for speculation what explains the change in 2008, possibly a significant change in the amount of money if it is as Browne and Cronin (2010) propose that money needs to be included in the commodity and inflation models. We believe a probable explanation for the shift is, that 2008 saw the end of a boom in commodity prices (as is seen from Figure 2), and as the boom ended so did the relation between inflation and commodities "normalize" as well. The prices of commodities had grown continuously from 1999 until crashing dramatically in the financial crisis of 2008. Likely some of the pre-2008 rally in prices was driven by speculation, and only partially rooted in reflecting the demand for commodities. Speculation can result in bubble if the average forecast of the speculators is incorrect (Suni and De Meo, 2009), and it is debatable that this was the case for commodities at the time.

The commodity prices have since recovered to some degree but the boom is conclusively over. The end of the commodity boom could improve their utility in inflation forecasts, since the prices will reflect more accurately changes in their global demand.

7 Conclusion

Our investigation centered on testing whether the inclusion of commodities in an ADL model, would improve Finnish inflation forecasts from those of a univariate AR model. The results were in line with the previous literature (mainly on US inflation), in that there is a link between commodity prices and inflation, but it is weak and not useful in improving the univariate forecasts.

Pseudo out-of-sample forecasts are rightly considered the gold standard for the evaluation of forecast models, and our results are compelling evidence that out-of-sample forecasts are absolutely necessary for credible assessment of forecasting models. Our results revealed that in-sample AIC was not a reliable predictor of out-of-sample forecast accuracy. The starkest case was the one month AR and MA models, where the MA forecast was greatly superior to the AR forecast despite its in-sample AIC being considerably higher. Similarly none of the commodity ADL models significantly improved (and many did worse) on the comparable AR model despite their better in-sample fits and lower AIC.

The discord between the in and out-of-sample results aroused mainly from structural change, but whether the problem would have been smaller with more parsimonious models specified by BIC is something to ponder. Simpler models could be more robust to changes in the forecasting environment, and thus less adversely affected by them. Obviously being approximately right is preferable to being accurately wrong, so maybe more thought on aiming for model simplicity is warranted.

The relation between commodities and inflation was not stable during the forecast period as evidenced by the changes in the long run means of the ADL models. The changes could have risen from many different factors: perhaps the manner in which the ECB reacts to commodity inflation has changed, or the link between commodity and consumer prices could have changed as well. Certainly as is seen from Figure 11 in 2009 the coefficients of the 1 month models experienced an abrupt and considerable change, which can be seen as evidence for a change in the inflation regime.

While it might be still possible to design a model successfully forecasting inflation from commodity prices, we are not convinced of it being a worthwhile pursuit. There simply are too many possible ways for any model of this kind breaking down. Inflation itself can have structural changes, as can change the way central banks react to inflationary pressures from rising commodity prices, technological developments can change the demand of commodities in manufacturing, to list just a few possible conundrums. Essentially there is not much reason to believe in the stability of the relation between commodities and inflation. While commodity prices should still serve a valuable purpose in modeling the causes of inflation, this does not generalize to them necessarily being useful in inflation forecasting.

In regard to inflation forecasting, since there exists an abundance of univariate models of every imaginable kind offering at least adequate inflation forecasts, it would seem sensible to concentrate the inflation forecasting efforts on univariate models. Since multivariate models do not significantly improve the forecasts, and unnecessarily introduce new ways on how the model can break down (changes in the relation between inflation and the exogenous variable), it is unclear why they should be preferred over univariate models.

Rather harshly Atkeson and Ohanian (2001) advised abandoning the search for Phillips-curve based inflation forecasting models altogether, and we wonder should the same conclusion be drawn also for commodity based forecasting models. We feel that abandoning the search in the case of commodities would be premature. While the commodity prices may not be particularly useful at the moment for inflation forecasting, and the topic has already been extensively researched, this does not yet hold for the variables derived from them. Especially models based on convenience yields seem promising avenues for the future of inflation forecasting (Gospodinov and Ng, 2010). Thus there still might be a role for commodities in inflation forecasting, but it will likely not be in models straightforwardly using untransformed prices, but rather in forecasts from convenience yields or some other imaginative variable to be derived from commodities.

8 References

Acharya, Ram N., Paul F. Gentle and Krishna P. Paudel (2010), "Examining the CRB index as a leading indicator for US inflation", Applied Economics Letters 17: 1493-1496.

Atkeson, Andrew, and Lee E. Ohanian (2001), "Are Phillips Curves Useful for Forecasting Inflation?", Federal Reserve Bank of Minneapolis Quarterly Review Volume 25, No. 1.

Armesto, Michael T., and William T. Gavin (2005), "Monetary Policy and Commodity Futures", Federal Reserve of St. Louis Review 87: 395-405.

Bloch, Harry, A. Michael Dockery and David Sapsford (2006), "Commodity Prices and the Dynamics of Inflation in Commodity-Exporting Nations: Evidence from Australia and Canada", The Economic Record Vol. 82. Special Issue: 97-109.

Blomberg, Brock S., and Ethan S. Harris (1995), "The Commodity-Consumer Price Connection: Fact or Fable?", Federal Reserve Bank of New York, Economic Policy Review, October.

Boughton, James M., and William H. Branson (1988), "Commodity Prices as a Leading Indicator of Inflation", NBER Working Paper Series 2750.

Browne, Frank, and David Cronin (2010), "Commodity prices, money and inflation", Journal of Economics and Business 62: 331-345.

Cecchetti, Stephen G., Rita S. Chu, and Charles Steindel (2000), "The Unreliability of Inflation Indicators", Federal Reserve Bank of New York, Current Issues in Economics and Finance Volume 6, Number 4.

Cecchetti, Stephen G., and Richhild Moessner (2008), "Commodity prices and inflation dynamics", BIS Quarterly Review, December.

Ciner, Cetin (2011), "Commodity prices and inflation: Testing in the frequency domain", Research in

International Business and Finance 25: 229-237.

Dow Jones Indexes (2010), “The Dow Jones-UBS Commodity Index Handbook, as of June 2010”.

Edelstein, Paul, (2007), “Commodity Prices, Inflation Forecasts, and Monetary Policy”, Working Paper, University of Michigan.

Ellison, Martin, Markku Lanne, Antti Ripatti and Pentti Saikkonen (2010), “Non-Causal Inflation”, mimeo.

Eugeni, Francesca and Joel Kruger (1994), “The Ups and Downs of Commodity Price Indexes”, Federal Reserve Bank of Chicago, Chicago Fed Letter Number 88.

Furlong, Fred and Robert Ingenito (1996), “Commodity Prices and Inflation”, Federal Reserve Bank of San Francisco Economic Review 1996, Number 2.

Garner, Alan C. (1995), “How Useful Are Leading Indicators of Inflation?”, Federal Reserve Bank of Kansas City Economic Review, Second Quarter.

Gospodinov, Nikolai, and Serena Ng (2010), “Commodity Prices, Convenience Yields and Inflation”, mimeo.

Hanson, Michael S. (2004), “The ‘price puzzle’ reconsidered”, Journal of Monetary Economics 51: 1385-1413.

Kwan, Simon (2005), “Inflation Expectations: How the Market Speaks”, Federal Reserve Bank of San Francisco Economic Letter Number 2005-25.

Kyrtsou, Catherine and Walter C. Labys (2006), “Evidence for chaotic dependence between US inflation and commodity prices”, Journal of Macroeconomics 28: 256-266.

Mahdavi, Saeid, and Su Zhou (1997), “Gold and Commodity Prices as Leading Indicators of Inflation: Tests of Long-Run Relationship and Predictive Performance”, Journal of Economics and Business 49:

475-489.

Malliaris, A.G. (2006), "US inflation and commodity prices: Analytical and empirical issues", Journal of Macroeconomics 28: 267-271.

Marcellino, Massimiliano, James H. Stock and Mark W. Watson (2005), "A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series", mimeo.

Moosa, Imad A. (1998), "Are Commodity Prices a Leading Indicator of Inflation?", Journal of Policy Modeling 20: 201-212.

Pindyck, Robert S. and Julio J. Rotemberg (1990), "The Excess Co-Movement of Commodity Prices", The Economic Journal 403: 1173-1189.

Stock, James H., and Mark W. Watson (2003), "Forecasting Output and Inflation: The Role of Asset Prices", Journal of Economic Literature 41: 788-829.

Stock, James H., and Mark W. Watson (2006), "Why Has U.S. Inflation Become Harder to Forecast?", NBER Working Paper Series 12324.

Stoll, Hans R., and Robert E. Whaley (2009), "Commodity Index Investing and Commodity Futures Prices", Working Paper, Vanderbilt University.

Suni, Paavo and Emanuele De Meo (2009), "Speculation and Commodity Prices", Suhdanne 2009:1.

Tsay, Ruey S. (2005), "Analysis of Financial Time Series, Second Edition", Wiley-Interscience.

Verbeek, Marno (2004), "A Guide to Modern Econometrics, Second Edition", John Wiley & Sons Ltd.

Webb, Roy H. (1988), "Commodity Prices as Predictors of Aggregate Price Change", Federal Reserve Bank of Richmond, Economic Review, November/December.

Appendix 1

Commodity Index Weights

	DJUBS	IMF	IMFNF
Energy	**32.4**	**63.2**	**0.0**
Crude	15.0	53.6	
Natural Gas	10.4	7.0	
Heating Oil	3.5		
Unleaded Gasoline	3.5		
Coal		2.5	
Industrial Metal	**17.0**	**10.7**	**28.9**
Copper	7.3	2.8	7.7
Aluminium	4.9	3.9	10.5
Zinc	2.5	0.6	1.7
Nickel	2.2	1.1	3.0
Iron, Tin, Lead, Uranium		2.2	6.0
Precious Metal	**14.7**	**0.0**	**0.0**
Gold	10.8		
Silver	3.9		
Agricultural Raw Materials	**1.8**	**7.7**	**20.8**
Cotton	1.8	0.7	1.8
Timber		3.4	9.1
Hides		2.6	7.1
Rubber		0.6	1.5
Wool		0.5	1.3
Meat and Seafood	**5.2**	**6.9**	**18.7**
Beef	3.3	1.4	3.9
Pork	1.9	1.1	3.1
Seafood		3.2	8.6
Poultry		0.9	2.4
Lamb		0.3	0.7
Foods	**26.5**	**9.8**	**26.5**
Wheat	4.8	1.7	4.5
Corn	7.1	1.0	2.8
Soybeans	5.5	1.2	3.3
Soybean Oil	2.7	0.4	1.2
Soy Meal	2.6	0.8	2.3
Sugar	3.9	0.9	2.4
Plant Oils		1.5	4.1
Rice		0.6	1.7
Barley		0.3	0.7
Bananas		0.4	1.1
Oranges		0.5	1.3
Groundnuts		0.2	0.6
Fishmeal		0.2	0.5
Beverages	**2.4**	**1.8**	**4.9**
Coffee	2.4	0.8	2.3
Tea		0.3	0.8
Cocoa Beans		0.7	1.8

Appendix 2

The ADL equation: $\Delta\pi_t = c + \sum\theta_i \Delta\pi_{t-i} + \sum\phi_j \Delta\log X_{t-j} + \varepsilon_t$

In the long-run $\Delta\pi_t \equiv \Delta\pi$ and $\Delta\log X_t \equiv \Delta\log X$ and $\varepsilon_t = 0$

$$\Delta\pi = c + \sum\theta_i \Delta\pi + \sum\phi_j \Delta\log X \Leftrightarrow$$

$$\Delta\pi(1 - \sum\theta_i) = c + \sum\phi_j \Delta\log X \Leftrightarrow$$

$$\Delta\pi = \frac{c}{1 - \sum\theta_i} + \frac{\sum\phi_j}{1 - \sum\theta_i} \Delta\log X$$