CrossMark

On inflation expectations in the NKPC model

Philip Hans Franses[1]

Abstract To create an estimable version for annual data of the hybrid new Keynesian Phillips curve, one needs an expression for the expectation of next year's inflation. The rational expectations literature assumes that this expectation is equal to the realization in the next year and an associated forecast error. This paper argues that this assumption goes against the Wold decomposition theorem, and that it introduces correlation between the error and a regressor. A more appropriate approach resorts to a MIDAS type of model, where forecast updates for next year are created when for example monthly inflation rates come in. An illustration to annual USA inflation, 1956–2016, shows the merits of this MIDAS approach.

1 Introduction

The so-called hybrid new Keynesian Phillips curve (NKPC) for inflation receives quite some attention in the recent literature. Mavroeidis et al. (2014) provide an excellent survey, and earlier substantive accounts can be found in Calvo (1983), Gali (2008) and Woodford (2003). A key feature in this literature is the inclusion of inflation expectations in empirical models. Mavroeidis et al. (2014) provide an overview of various ways to include those expectations, which range from the inclusion of various

✉ Philip Hans Franses
franses@ese.eur.nl

1 Econometric Institute, Erasmus School of Economics, Burgemeester Oudlaan 50,
3062 PA Rotterdam, The Netherlands

observable variables to survey-based expectations, see for example Preston (2005) for an interesting account on surveys data. To set matters more precise: when considering annual data, a basic version of a NKPC model for this year's inflation includes as explanatory variables the one-year lagged inflation, a measure of marginal costs, and the expected value for next year's inflation. The present paper focuses on this last variable in this basic type of model.

In one part of the literature on the NKPC, notably Lanne and Luoto (2013), Gali and Gertler (1999), and Gali et al. (2005), among others, a key assumption is rational expectations. There it is commonly assumed that the expected value of next year's inflation is equal to the realized value of inflation in the next year plus the forecast error for next year. Substituting the latter two terms in the model renders a model for this year's inflation being a function of one-year-ahead and one-year lagged inflation, and the measure for marginal costs. Abstaining from the latter variable, in time series terms, this model for inflation provides an exciting opportunity to include the past and the future.

In the present paper, I argue that this assumption made in the RE literature is problematic for at least two reasons. First, the well-known Wold decomposition theorem says that any time series is decomposable into a deterministic component (like a constant and a trend) and a weighted sum of all current and past shocks (or prediction errors). More precise for the case at hand, it is thus impossible that a current observation on inflation depends on a shock to inflation in the future. Of course, if an event will take place, with known impact on inflation (think of a devaluation of a currency), it can be included. Note that it is then not a shock but a deterministic event. Expert adjustment of forecasts typically involves such actions. The second reason why the assumption in the RE literature is problematic is that by construction the error term in the estimable version of the NKPC model is perfectly correlated with one of the regressors. Indeed, again Wold's decomposition theorem tells us that the expected value of the covariance between the current observation and the current shock is equal to the variance of the shocks. To "instrument away" this nonzero covariance may be a difficult task.

In this paper, I propose an alternative approach to including the expected value of inflation. One could perhaps think of including this year's inflation rate or past year's inflation rate, but this leads to the un-identifiability of the parameter associated with the expectations. It seems better to resort to a MIDAS-type approach, which here entails that the forecast for next year's inflation is created based on information that comes in as the current year proceeds. That information can be based on all kinds of monthly, or even weekly and daily sources. Think, for example, of daily observable hotel prices via booking sites, or flight ticket prices. Just as an illustration, I will simply use monthly inflation figures, inspired by Frijns and Margaritis (2008) who use early-in-the-day volatility estimates to predict overall daily volatility. The MIDAS approach means that each month, one can update the NKPC model and re-estimate its parameters, because new information comes in.

The outline of this letter is as follows. Section 2 formalizes the notions above using simple expressions. Section 3 provides an illustration for USA inflation data and shows, for example, that only using the first 4 months of the year with inflation data delivers already quite accurate predictions. Section 4 concludes.

2 An analysis

This section shows that replacing the expected value of next year's inflation by the realized value of inflation in the next year plus the forecast error for next year might not be a good idea. Next, I provide alternative approaches, where the one based on MIDAS-type modeling seems most useful.

Denote π_t as the annual inflation rate, x_t as a measure of marginal costs, and $E_t\pi_{t+1}$ as the one-year-ahead expected value of inflation made at time t. The hybrid new Keynesian Phillips curve (NKPC) can be summarized as

$$\pi_t = \mu + \alpha E_t\pi_{t+1} + \beta\pi_{t-1} + \gamma x_t \tag{1}$$

See equation (4) in Lanne and Luoto (2013). To estimate the parameters in this NKPC model, one needs to replace the unobserved variable $E_t\pi_{t+1}$ by an observable variable. In the rational expectations (RE) literature, it is custom to assume that

$$E_t\pi_{t+1} = \pi_{t+1} + \omega_{t+1}$$

See Lanne and Luoto (2013, p. 564). This results in an estimable version of the NKPC model like

$$\pi_t = \mu + \alpha\pi_{t+1} + \beta\pi_{t-1} + \varepsilon_t \tag{2}$$

with

$$\varepsilon_t = \alpha\omega_{t+1} + \gamma x_t$$

The key assumption in the RE literature is that $\alpha\omega_{t+1}$ is distributed as independently and identically (IID) over time. With this assumption, Lanne and Luoto (2013) use Maximum Likelihood to estimate the parameters in (2).

There is, however, a problem with this approach. The IID assumption of $\alpha\omega_{t+1}$ may perhaps be defendable, but the key issue is that the error term in (2) is not independent from one of the regressors. In fact,

$$E(\pi_{t+1}, \varepsilon_t) = \alpha E(\pi_{t+1}\omega_{t+1}) + \gamma E(\pi_{t+1}x_t)$$

First, one may wonder whether $E(\pi_{t+1}x_t)$ is equal to 0, and that seems hard to verify using empirical data. Certainly, it holds that

$$E(\pi_{t+1}\omega_{t+1}) = E(E_t\pi_{t+1} - \omega_{t+1}, \omega_{t+1}) = -\sigma_\omega^2 \neq 0$$

This makes a regressor and the error term in (2) to be correlated.

This insight basically follows from the familiar Wold decomposition theorem. This theorem says that any time series y_t can be written as the sum of a deterministic component μ_t, including, for example, a constant and a trend, and a component that includes a weighted average of current and past shocks ε_t, that is,

$$y_t = \mu_t + \theta_0\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots = \mu_t + \sum_{i=0}^{\infty}\theta_i\varepsilon_{t-i}$$

Usually, θ_0 is set equal to 1. Given availability of past shocks and the parameters for these past shocks and the deterministic terms, it is clear that the Wold decomposition implies that

$$E_t y_{t+1} = \mu_{t+1} + \sum_{i=1}^{\infty}\theta_i\varepsilon_{t+1-i}$$

Hence, as μ_{t+1} is deterministic and perfectly forecastable,

$$y_{t+1} = E_t y_{t+1} + \varepsilon_{t+1}$$

In words, a time series is the sum of a predictable part and an unpredictable part, where the latter is also called the forecast error. This expression also shows that

$$E\left(y_{t+1}\varepsilon_{t+1}\right) = E\left(E_t y_{t+1} + \varepsilon_{t+1}, \varepsilon_{t+1}\right) = 0 + E\left(\varepsilon_{t+1}^2\right) = \sigma_\varepsilon^2$$

So, replacing $E_t\pi_{t+1}$ in (1) by

$$E_t\pi_{t+1} = \pi_{t+1} + \omega_{t+1}$$

does not seem the best option.

What then could we do? Let us go back to

$$\pi_t = \mu + \alpha E_t\pi_{t+1} + \beta\pi_{t-1} + \gamma x_t$$

One may now decide to replace $E_t\pi_{t+1}$ by π_t. This is also not a good idea for two reasons. First, during a year, there is no information on π_t in that total year, and only at the end of the year, we know this year's annual inflation rate. The second reason is that then (1) becomes

$$\pi_t = \mu + \alpha\pi_t + \beta\pi_{t-1} + \gamma x_t$$

This is equivalent to

$$\pi_t = \frac{\mu}{1-\alpha} + \frac{\beta}{1-\alpha}\pi_{t-1} + \frac{\gamma}{1-\alpha}x_t$$

In that case, the parameters in the NKPC model are not identifiable.

The same problem arises when one decides to replace $E_t\pi_{t+1}$ by π_{t-1}. This leads to

$$\pi_t = \mu + (\alpha + \beta)\pi_{t-1} + \gamma x_t$$

where again the parameters are not identified.

3 A MIDAS-based solution

There is a simple solution though, and that is that you do not replace $E_t \pi_{t+1}$ by π_t, but, for example, by the monthly inflation rates that come in as the year proceeds. These models are called MIDAS models, see Ghysels et al. (2006, 2007), Foroni et al. (2015), and Breitung and Roling (2015) among many possible references.

Denoting $\pi_{s,t}$ as the inflation rate in month s of year t, relative to the same month s in the previous year $t-1$, that is,

$$\pi_{s,t} = 100(\log \mathrm{CPI}_{s,t} - \log \mathrm{CPI}_{s,t-1})$$

where $\mathrm{CPI}_{s,t}$ is the consumer price index in month s of year t. One can now replace $\alpha E_t \pi_{t+1}$ by

$$\alpha_1 \pi_{1,t}$$
$$\alpha_1 \pi_{1,t} + \alpha_2 \pi_{2,t}$$
$$\alpha_1 \pi_{1,t} + \alpha_2 \pi_{2,t} + \alpha_3 \pi_{3,t}$$

and so on, until

$$\alpha_1 \pi_{1,t} + \alpha_2 \pi_{2,t} + \alpha_3 \pi_{3,t} + \cdots + \alpha_{12} \pi_{12,t}$$

In words, this says that the forecast for next year's inflation is first based on the inflation rate in January, and next it is based on the inflation rates in January and February of the current year, and so on. Note that the average of the twelve $\pi_{s,t}$ terms is not equal to π_t as

$$100 \left(\log \mathrm{CPI}_t - \log \mathrm{CPI}_{t-1}\right) \neq \frac{1}{12} \sum_{s=1}^{12} 100(\log \mathrm{CPI}_{s,t} - \log \mathrm{CPI}_{s,t-1})$$

Hence, during the current year, one can use monthly data as input to forecasts for next year's inflation.

4 Illustration for US annual inflation rates

Figure 1 displays the annual inflation rates for the USA, for the sample 1956–2016. Figure 2 displays the annualized inflation rate observed in January as well as the yearly data. Clearly, there is substantial common variation in the data. Figure 3 contains all the monthly inflation rates and shows that there can be sizable variation in the data within a year.
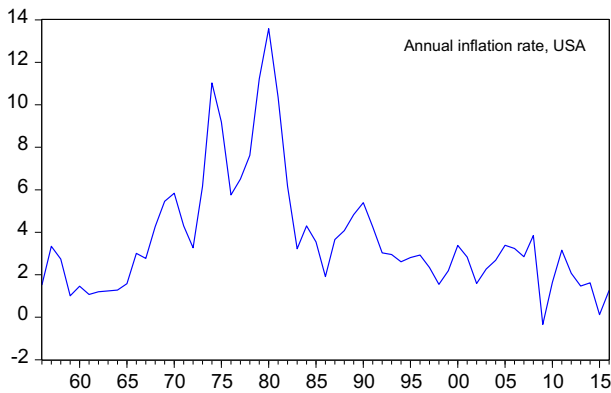
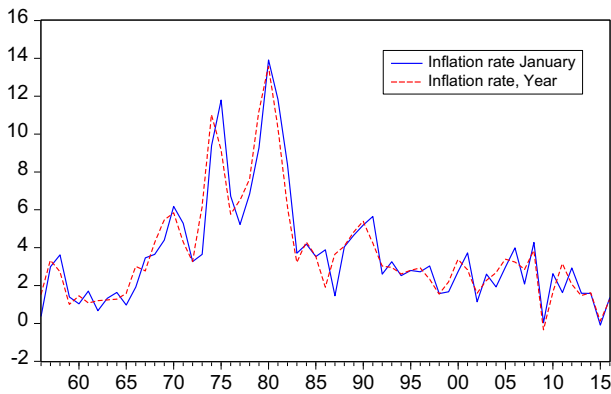**Fig. 1** Annual CPI-based inflation, USA, 1956–2016

Fig. 2 Annual inflation rate versus the inflation rate in January of the same year

There are various MIDAS-type models to consider. The first type assumes that each month new and relevant information might come is, and these models are

$$\pi_t = \mu + \beta\pi_{t-1} + \varepsilon_t$$
$$\pi_t = \mu + \alpha_1\pi_{1,t} + \beta\pi_{t-1} + \varepsilon_t$$
$$\pi_t = \mu + \alpha_1\pi_{1,t} + \alpha_2\pi_{2,t} + \beta\pi_{t-1} + \varepsilon_t$$
$$\cdots$$
$$\pi_t = \mu + \alpha_1\pi_{1,t} + \alpha_2\pi_{2,t} + \cdots + \alpha_{12}\pi_{12,t} + \beta\pi_{t-1} + \varepsilon_t$$

The parameters are estimated unrestrictedly, thereby following the format recommended in Foroni et al. (2015), which is called the unrestricted MIDAS model, or in short, UMIDAS.
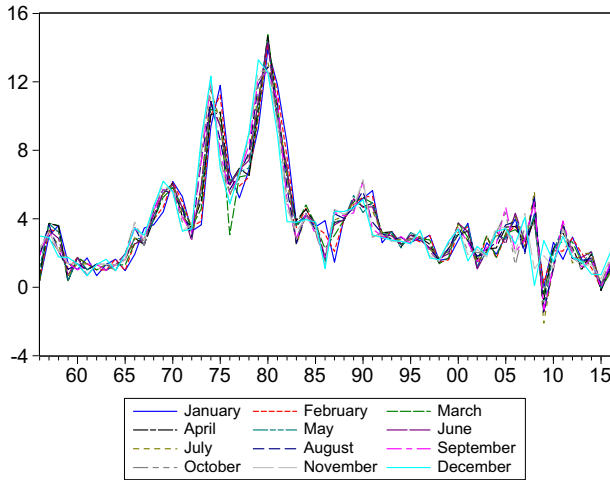
Fig. 3 Annualized inflation rates per month

Table 1 gives a selection of the estimation results.[1] It is clear that there is substantial variation across the estimated α_s parameters across the UMIDAS models. This reinforces that imposing structure on the parameters using, for example, Almon lags, as is often done in the literature, does not make sense here.

Figure 4 presents the root-mean-squared prediction error and mean absolute error for the test sample 2005–2016, where the parameters are estimated for 1956–2004. The UMIDAS model includes January (1), January and February (2), …, January–December (12). The graphs in Fig. 4 show that already quite some accurate forecasts can be obtained when the model

$$\pi_t = \mu + \alpha_1 \pi_{1,t} + \alpha_2 \pi_{2,t} + \alpha_3 \pi_{3,t} + \alpha_4 \pi_{4,t} + \beta \pi_{t-1} + \varepsilon_t$$

is considered. This can also be learned from Fig. 5 which shows a sharp increase in the R^2 when the months January–April are included. Figure 6 gives a graphical impression of how the parameter estimates develop.

Now that UMIDAS models are considered, it is also possible to look at alternative forecast schedules, like, for example,

$$\pi_t = \mu + \alpha_1 \pi_{1,t} + \alpha_2 \pi_{2,t} + \cdots + \alpha_{12} \pi_{12,t} + \beta \pi_{t-1} + \varepsilon_t$$
$$\pi_t = \mu + \alpha_2 \pi_{2,t} + \cdots + \alpha_{12} \pi_{12,t} + \beta \pi_{t-1} + \varepsilon_t$$
$$\pi_t = \mu + \alpha_3 \pi_{3,t} + \cdots + \alpha_{12} \pi_{12,t} + \beta \pi_{t-1} + \varepsilon_t$$
$$\cdots$$

---

[1] When possible, given the degrees of freedom, all estimated models were examined using the Quandt-Andrews unknown breakpoint test, where the null hypothesis is "No breakpoints with 15% trimmed data". For most estimated models this null hypothesis is not rejected. Detailed estimation results are available upon request.

Table 1 Estimation results for the models

$$\pi_t = \mu + \beta \pi_{t-1} + \varepsilon_t$$

$$\pi_t = \mu + \alpha_1 \pi_{1,t} + \beta \pi_{t-1} + \varepsilon_t$$

$$\pi_t = \mu + \alpha_1 \pi_{1,t} + \alpha_2 \pi_{2,t} + \beta \pi_{t-1} + \varepsilon_t$$

$$\pi_t = \mu + \alpha_1 \pi_{1,t} + \alpha_2 \pi_{2,t} + \cdots + \alpha_{12} \pi_{12,t} + \beta \pi_{t-1} + \varepsilon_t$$

β	*0.187 (0.076)*	-0.192 (0.103)	-0.084 (0.071)	-0.002 (0.003)
α_1		*1.077 (0.099)*	-0.817 *(0.243)*	*0.080 (0.010)*
α_2			*1.835 (0.226)*	*0.123 (0.013)*
α_3				0.015 (0.008)
α_4				*0.123 (0.015)*
α_5				*0.062 (0.019)*
α_6				*0.105 (0.014)*
α_7				*0.066 (0.012)*
α_8				*0.089 (0.012)*
α_9				*0.089 (0.015)*
α_{10}				*0.077 (0.015)*
α_{11}				*0.108 (0.015)*
α_{12}				*0.066 (0.009)*
R^2	0.665	0.891	0.950	1.000[a]

Bold and italic means significant at the 5% level
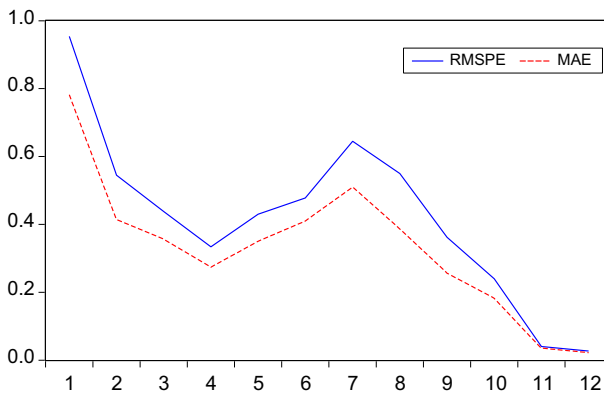
[a] The true score is 0.99959

Fig. 4 Root-mean-squared prediction error and mean absolute error for the test sample 2005–2016, where the parameters are estimated for 1956–2004. The UMIDAS model includes January (1), January and February (2), …, January–December (12)
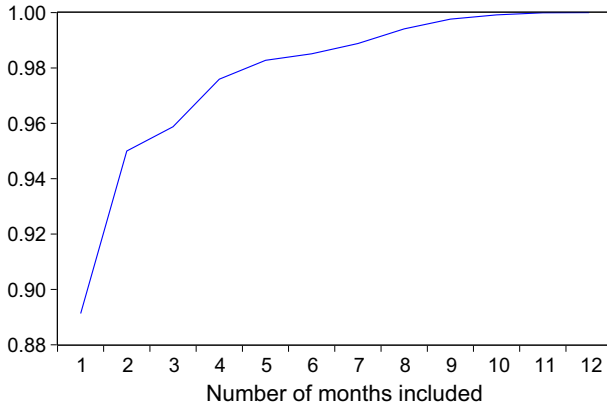
Fig. 5 The R^2 of the UMIDAS models where each time an additional month is included
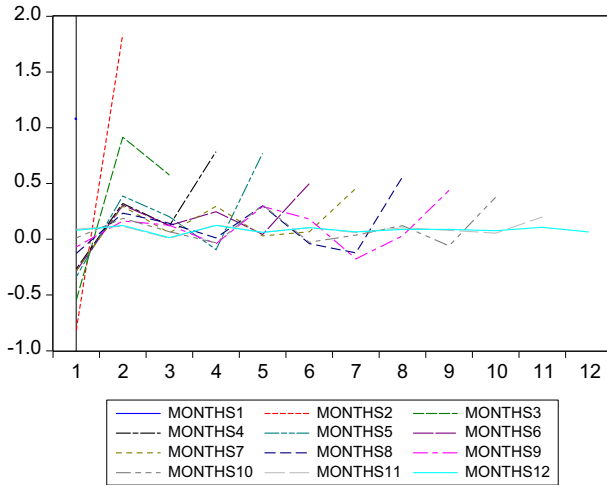

Fig. 6 Estimated parameters in UMIDAS models, where each time an additional month is included

$$\pi_t = \mu + \alpha_{12}\pi_{12,t} + \beta\pi_{t-1} + \varepsilon_t$$
$$\pi_t = \mu + \beta\pi_{t-1} + \varepsilon_t$$

A selection of estimation results is presented in Table 2. Looking at the significant parameters, it is clear that there is wide variety of possible relevant models.

Finally, it is also possible to see which of the monthly inflation rates can be viewed as the most informative for forecasting next year's inflation. For that purpose, one can consider

$$\pi_t = \mu + \alpha_1\pi_{1,t} + \beta\pi_{t-1} + \varepsilon_t$$
$$\pi_t = \mu + \alpha_2\pi_{2,t} + \beta\pi_{t-1} + \varepsilon_t$$
$$\cdots$$
$$\pi_t = \mu + \alpha_{12}\pi_{12,t} + \beta\pi_{t-1} + \varepsilon_t$$

Table 2 Estimation results for the models

$$\pi_t = \mu + \alpha_1\pi_{1,t} + \alpha_2\pi_{2,t} + \cdots + \alpha_{12}\pi_{12,t} + \beta\pi_{t-1} + \varepsilon_t$$

$$\pi_t = \mu + \alpha_2\pi_{2,t} + \cdots + \alpha_{12}\pi_{12,t} + \beta\pi_{t-1} + \varepsilon_t$$

$$\pi_t = \mu + \alpha_3\pi_{3,t} + \cdots + \alpha_{12}\pi_{12,t} + \beta\pi_{t-1} + \varepsilon_t$$

$$\pi_t = \mu + \alpha_{12}\pi_{12,t} + \beta\pi_{t-1} + \varepsilon_t$$

β	−0.002 (0.003)	0.006 (0.004)	*0.060 (0.008)*	*0.337 (0.051)*
α_1	*0.080 (0.010)*			
α_2	*0.123 (0.013)*	*0.210 (0.012)*		
α_3	0.015 (0.008)	0.016 (0.012)	*0.131 (0.026)*	
α_4	*0.123 (0.015)*	*0.093 (0.023)*	*0.143 (0.061)*	
α_5	*0.062 (0.019)*	*0.095 (0.029)*	0.074 (0.078)	
α_6	*0.105 (0.014)*	*0.094 (0.022)*	*0.153 (0.058)*	
α_7	*0.066 (0.012)*	*0.070 (0.018)*	0.062 (0.049)	
α_8	*0.089 (0.012)*	*0.079 (0.018)*	0.014 (0.047)	
α_9	*0.089 (0.015)*	*0.094 (0.023)*	*0.203 (0.060)*	
α_{10}	*0.077 (0.015)*	*0.093 (0.023)*	0.038 (0.061)	
α_{11}	*0.108 (0.015)*	*0.073 (0.022)*	0.029 (0.059)	
α_{12}	*0.066 (0.009)*	*0.078 (0.014)*	*0.100 (0.038)*	*0.676 (0.050)*
R^2	1.000[a]	1.000[b]	0.999	0.921

Bold and italic means significant at the 5% level
[a]The true score is 0.999959
[b]The true score is 0.999896

The estimation results appear in Table 3. The peak R^2 value appears in August (0.983). On the other hand, around April and May the R^2 values are already quite high, and also the associated parameter is close to 1, with a small standard error. In other words, predictions from the NKPC model based on data available to and including April are rather accurate for the finally to be obtained annual end-of-the-year inflation.[2]

---

[2] Upon suggestion of one the reviewers, the MIDAS models were extended with survey forecasts. The Michigan survey data were obtained from https://fred.stlouisfed.org/series/MICH. The data start in January 1978. They are available monthly. Each time the respondents are asked to make a forecast for the next 12 months. For the present purposes this means that only the quote in December each year is useful. When this variable is added to the model, the associated parameter is not significant. Next, the data from the Survey of Professional Forecasters are obtained from https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters. These expectations for next year's inflation are collected every quarter since 1981Q3. Hence, these expectations are the same within a particular quarter. The inclusion of these survey-based forecasts results in four new MIDAS models. Wald tests and t tests on the significant of these variables all resulted in the p values much higher than 5%. Details on the computations are available upon request.

Table 3 Estimation results for

$$\pi_t = \mu + \alpha_1\pi_{1,t} + \beta\pi_{t-1} + \varepsilon_t$$
$$\pi_t = \mu + \alpha_2\pi_{2,t} + \beta\pi_{t-1} + \varepsilon_t$$
$$\cdots\cdots$$
$$\pi_t = \mu + \alpha_{12}\pi_{12,t} + \beta\pi_{t-1} + \varepsilon_t$$

Month	α_s	β	R^2
January	*1.077 (0.099)*	-0.192 (0.103)	0.891
February	*1.104 (0.068)*	*-0.186 (0.070)*	0.940
March	*0.918 (0.049)*	0.060 (0.050)	0.953
April	*0.976 (0.037)*	-0.003 (0.038)	0.974
May	*0.934 (0.032)*	0.045 (0.033)	0.979
June	*0.906 (0.032)*	*0.067 (0.033)*	0.978
July	*0.873 (0.029)*	*0.096 (0.030)*	0.980
August	*0.863 (0.027)*	*0.132 (0.027)*	0.983
September	*0.806 (0.027)*	*0.190 (0.028)*	0.980
October	*0.791 (0.033)*	*0.219 (0.034)*	0.970
November	*0.739 (0.043)*	*0.272 (0.045)*	0.945
December	*0.676 (0.050)*	*0.337 (0.051)*	0.921

Bold and italic means significant at the 5% level

5 Conclusion

The hybrid new Keynesian Phillips curve model for annual inflation involves an expectation of next year's inflation. The common assumption in the rational expectations literature is to include the actual next year's inflation and prediction error. This assumption leads to two inconveniences, that is, endogeneity of one of the regressors, and it violates the Wold decomposition theorem. A simple solution is presented which relies on the MIDAS notion, that is, higher frequency data within the same year can be used to create forecasts for next year's inflation. An illustration to US annual inflation rates with incoming monthly inflation rates showed the merits of this approach.

Compliance with ethical standards

Conflict of interest The author declares that there is no conflict of interest.

References

Breitung J, Roling C (2015) Forecast inflation rates using daily data: a nonparametric MIDAS approach. J Forecast 34:588–603

Calvo G (1983) Staggered prices in a utility-maximizing framework. J Monet Econ 12:383–398

Foroni C, Marcellino M, Schumacher C (2015) Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. J R Stat Soc Ser A 178:57–82

Frijns B, Margaritis D (2008) Forecasting daily volatility with intraday data. European Journal of Finance 14:523–540

Gali J (2008) Monetary policy, inflation, and the business cycle: an introduction to the New Keynesian framework. Princeton University Press, Princeton

Gali J, Gertler M (1999) Inflation dynamics: a structural econometric approach. J Monet Econ 44:195–222

Gali J, Gertler M, Lopez-Salido JD (2005) Robustness of the estimates of the hybrid New Keynesian Phillips curve. J Monet Econ 52:1107–1118

Ghysels E, Santa-Clara P, Valkanov R (2006) Predicting volatility: getting the most out of return data sampled at different frequencies. J Econom 131:59–95

Ghysels E, Sinko A, Valkanov R (2007) MIDAS regressions: further results and new directions. Econom Rev 26:53–90

Lanne M, Luoto J (2013) Autoregression-based estimation of the new Keynesian Phillips curve. J Econ Dyn Control 37:561–570

Mavroeidis S, Plagborg-Møller M, Stock JH (2014) Empirical evidence of inflation expectations in the new Keynesian Phillips curve. J Econ Lit 52:124–188

Preston B (2005) Learning about monetary policy rules when long-horizon expectations matter. Int J Cent Bank 1:81–126

Woodford M (2003) Interest rates and prices: foundations of a theory of monetary policy. Princeton University Press, Princeton