Inflation Dynamics and the Hybrid Neo Keynesian Phillips Curve: The Case of Chile

Medel, Carlos

University of Nottingham, United Kingdom

6 March 2015

Inflation Dynamics and the Hybrid Neo Keynesian Phillips Curve: The Case of Chile

Carlos A. Medel*
School of Economics
University of Nottingham, United Kingdom

March 6, 2015

Abstract

It is recognised that the understanding and accurate forecasts of key macroeconomic variables are fundamental for the success of any economic policy. In the case of monetary policy, many efforts have been made towards understanding the relationship between past and expected values of inflation, resulting in the so-called Hybrid Neo-Keynesian Phillips Curve (HNKPC). In this article I investigate to which extent the HNKPC help to explain inflation dynamics as well as its out-of-sample forecast, for the case of the Chilean economy. The results show that the forward-looking component is significative and accounts from 1.58 to 0.40 times the lagged inflation coefficient. Also, I find predictive gains close to 45% (respect to a backward-looking specification) and up to 80% (respect to the random walk) when forecasting at 12-months ahead.

JEL-Codes: *C22*, *C53*, *E31*, *E37*, *E47*.

Keywords: *New Keynesian Phillips Curve*, *inflation forecast*, *out-of-sample comparisons*, *survey data*, *real-time dataset*.

Abstract

Es ampliamente reconocido que la comprensión y precisión de los pronósticos de las principales variables macroeconómicas son fundamentales para el éxito de cualquier política económica. En el caso de la política monetaria, muchos esfuerzos han sido realizados para la comprensión de la relación entre valores esperados y rezagados de la inflación, resultando en la llamada Curva de Phillips Híbrida Neokeynesiana (HNKPC). En este artículo se investiga en qué medida la HNKPC ayuda a explicar la dinámica inflacionaria, así como su pronóstico fuera de muestra, para el caso de la economía chilena. Los resultados muestran que el coeficiente de expectativas es significativo y representa desde 1,58 hasta 0,40 veces el coeficiente de la inflación rezagada. demás, se encuentran ganancias predictivas cercanas al 45% (respecto a una especificación basada exclusivamente en rezagos) y de hasta un 80% (respecto a la caminata aleatoria) pronosticando 12 meses adelante.

Códigos JEL: *C22*, *C53*, *E31*, *E37*, *E47*.

Palabras clave: *Curva de Phillips Neokeynesiana*, *proyección de inflación*, *comparación de pronósticos*, *encuestas*, *datos en tiempo real*.

*Address: 118 Talbot Street, Nottingham NG1 5GP, England. E-mail: lexcm6@nottingham.ac.uk.

1 Introduction

The aim of this article is to investigate to which extent forward-looking (FL) measures of inflation help to explain inflation dynamics as well as its out-of-sample behaviour with a Phillips Curve ensemble. This objective is tackled by analysing the performance of the so-called Hybrid Neo-Keynesian Phillips Curve (HNKPC), introduced by Galí and Gertler (1999, GG), using a dataset of the Chilean economy.

It is widely recognised that the understanding and accurate forecasts of key macroeconomic variables are fundamental for the success in almost all economic policies. In the case of monetary policy, inflation forecasts are not useful from a practical but from a theoretical viewpoint also. Many efforts have been made towards understanding the relationship between past and expected values of inflation. The former reflects the traditional inertia of price setting, while the latter stands as an ingredient of rational expectations agents' behaviour. The HNKPC offers an amalgamation of these two components by allowing both a Calvo price setting scheme plus a fraction of FL price-setters firms (see Calvo, 1983, and GG).

Suppose a staggered price-setting scheme. Let $1-\theta$ the fraction of firms that change prices at a given period, and $1-\omega$ the fraction of firms that set prices optimally in a FL manner. Hence, current prices constitute a weighted average between backward- (BL) and FL firms, leading to the HNKPC baseline equation:

$$\pi_t = \lambda x_t + \gamma_b \pi_{t-1} + \gamma_f \mathbb{E}_t[\pi^f_{t,t+h}] + \varepsilon_t, \tag{1}$$

where π_t is inflation, $\mathbb{E}_t[\pi^f_{t,t+h}]$ is the inflation expectation at period f, measured with a forecast made h-step ahead at period t, and x_t is a real marginal cost measure. $\{\lambda; \gamma_b; \gamma_f; \sigma^2_\varepsilon\}$ are parameters to be estimated, and ε_t is a cost-push shock, $\varepsilon_t \sim iid\mathcal{N}(0, \sigma^2_\varepsilon)$. This specification constitutes a reduced form of a structural NKPC with $\gamma_f = \beta\theta/\phi$, $\gamma_b = \omega/\phi$, $\lambda = [(1-\omega)(1-\theta)(1-\beta\theta)]/\phi$, where β is a discount rate, and $\phi = \theta + \omega[1-\theta(1-\beta)]$. Equation (1) results in a convenient form as it allows many price setting schemes, making possible simple forecasting exercises.[1]

Many of the empirical evidence of the HNKPC have been collected for industrialised economies. Some selected examples are Roberts (1997), GG, Galí, Gertler, and López-Salido (2005), Rudd and Whelan (2005, RW), and Brissimis and Magginas (2008) for the US, and Jean-Baptiste (2012) for the UK. The main difference in their methodology concerns inflation expectation proxies, real-time estimates with different data vintages, and the measurement of marginal costs.[2] A current controversial methodological discussion confronts the results obtained by RW in opposition to those of GG. While the former finds that lagged inflation is the major driver of current inflation, the latter states that is the FL component. This bifurcation is due to different specifications and estimation method assumptions. This article follows more closely the GG derivation of the HNKPC, with some minor twists.

More evidence on the HNKPC is provided by Paloviita and Mayes (2005) for a panel of OECD countries. The authors find an influential role for the expectations, but also they unveil the

[1] Literature regarding a formal theoretical derivation of the HNKPC can be found in Smets and Wouters (2003, 2005), Christiano, Eichembaum, and Evans (2005), Erceg and Levin (2003), and Collard and Dellas (2004), among others.

[2] It is worth mentioning that the US economy has richer conclusions on this matter as it has several sources of survey expectations data with a long sample span, as is the case of the Survey of Professional Forecasters (SPF) of the Federal Reserve Bank of Philadelphia, the Livingstone Survey, the Michigan Survey, the Greenbook, Consensus Forecasts, the Congressional Budget Office, and the Real-Time Data Set for Macroeconomists (Croushore and Stark, 2001).

controversial role of the output gap as a measure of marginal costs. Finally, for the case of Chile, little research has been conducted in this matter. Some exceptions are Céspedes, Ochoa, and Soto (2007) and Pincheira and Rubio (2010). The first article derives a NKPC from a structural microfounded model, and analyse their in-sample ability to explain inflation dynamics. The second article addresses the issue of the weak predictive power of purely BL PC with real-time data.

In this article I first estimate an unrestricted version of the HNKPC with Chilean data, to then compare its predictive power with a BL PC and traditional benchmarks predicting at h-months-ahead, $h = \{1; 3; 6; 12\}$. The dataset corresponds to monthly inflation, a monthly index of economic activity, and the expectations of the Chilean Survey of Professional Forecasters (ChSPF). The estimation is made through the Generalised Method of Moments (GMM). I make use of the so-called *core inflation* measure to conduct a robustness analysis. Such analysis is complemented with some recursive estimations to shed some light about parameter uncertainty and stability.

The results show that the FL inflationary component is statistically significative when is included in the specification. In size, accounts from 1.58 to 0.40 times the lagged inflation coefficient. Real-time ChSPF forecasts of output are also useful but as instruments.[3] When considering short-term forecasting, I find predictive gains close to 45% (respect to the BL specification) and up to 80% (respect to the random walk) when forecasting at 12-months-ahead. However, these gains are not statistically significant according to the traditional Giacomini and White (2006; GW) test. In sum, these results should be read carefully and just as a valid benchmark. The in-sample results for core inflation support the existence of the HNKPC. Nevertheless, predictive results suggest that core could be a process with higher memory. The output gap plays a key role delivering better results than similar benchmark.

The article proceeds as follows. In Section 2 I detail the econometric procedure, alongside the dataset utilised emphasising the output gap construction–an unobservable variable. Section 3 presents the empirical results divided in those obtained in-sample and those when predicting both measures of inflation. Finally, Section 4 concludes.

2 Econometric setup

The baseline specification is the Equation (1). To avoid part of the simultaneity in the variables of the RHS, I estimate Equation (1) with GMM. However, this method eliminates *methodological* simultaneity only, as the series exhibits a high correlation given their underlying data generating process. I make use of lagged observations of the variables as instruments (IV). Recall that the problem that GMM addresses is the orthogonality condition $\mathbb{E}_t[\mathbf{x}_t'\varepsilon_t]$ that no longer holds. Hence, it is needed to instrumentalise the $\mathbf{x}_t'$ matrix with another one, say $\mathbf{z}_t$, containing ℓ IV ($\ell \geq k$) which fulfils:

$$\mathbb{E}_{t-1}[(\pi_t - \lambda x_t + \gamma_b \pi_{t-1} + \gamma_f \mathbb{E}_t[\pi^f_{t,t+h}]) \times \mathbf{z}_{t-1}] = 0. \tag{2}$$

In this context, a formal test for IVs' suitability is analysed through the Hansen's J-statistic:

$$J(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{w}}_T) = \frac{1}{T}(\pi_t - \mathbf{x}_t'\widehat{\boldsymbol{\beta}})'\mathbf{z}_t \widehat{\mathbf{w}}_T^{-1} \mathbf{z}_t'(\pi_t - \mathbf{x}_t'\widehat{\boldsymbol{\beta}}), \tag{3}$$

where $\widehat{\mathbf{w}}_T$ is a $\ell \times \ell$ symmetric and positive-definite *weighting matrix*, as it weight the moments considered in the estimations. Hence, GMM finds the vector of coefficients:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{x}'\mathbf{z}\widehat{\mathbf{w}}_T^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}\widehat{\mathbf{w}}_T^{-1}\mathbf{z}'\mathbf{y}, \tag{4}$$

[3] This finding is in line with those of Orphanides and van Norden (2002, 2005) obtained for the US economy.

that minimises Equation (3). As $J(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{w}}_T) \sim \chi^2_{\ell-k}$, along with the estimated coefficients it is also reported the p-value that test the Null Hypothesis: $\mathbb{E}_T[J(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{w}}_T)] = 0$. If p-value $> \alpha$, the IV are valid at the α-level of significance.

The estimation of the weighting matrix is made according to Hansen (1982) recommendation–the inverse of covariance matrix, *i.e.* $\widehat{\mathbf{w}}_T = \widehat{\mathbf{s}}^{-1}$, and avoiding potential autocorrelation with the Newey-West HAC method. The estimation of both covariance matrices–for the two stages: IV and final regression–is set in the same manner. The whitening lag specification is set automatic, to be selected according the Bayesian Information Criterion (BIC) choosing in a maximum of 3 lags (following the rule $T^{1/3}$).

2.1 Data

Equation (1) involves three different kinds of series: actual inflation, inflation expectations, and the output gap. The source of all variables is the Central Bank of Chile (CBC). The available sample spans from 2000.1 to 2013.12 (168 observations). When forecasting, it is used the firsts 77 observations (2000.1-2006.5) as *estimation sample*, leaving the remaining 91 observations to *evaluation sample* (2006.6-2013.12). This scheme delivers 91 out-of-sample observations when predicting 1-step ahead, 89 for 3-, 86 for 6-, and 80 for 12-months ahead.

Actual inflation–*headline inflation*–corresponds to annual percentage change of the total CPI (index level, 2013=100), the same measuring units in which the inflation target is set. For robustness exercises, I make use of another inflation measure, the so-called *core inflation*. This corresponds to the CPI inflation but extracting the components of *Food and beverages* and *Energy* (reducing exogenous volatility).

The inflation expectations are provided by the ChSPF.[4] The ChSPF is informed at the beginning of each month. Inflation forecasts are delivered for 1-, 12-, and 24-months ahead, along with projections of GDP for the current and following year. It collects answers from academics, consultants, executives and private sector consultants who also report forecasts for other variables. Since each individual analyst's projections are not revealed, the median forecast is used. The ChSPF starts in 2000 and several times has changed its content. Except for minor changes made since 2004.11, it has remained unaltered. On average over the period 2000-2009, 35 analysts completed the questionnaire each month.

Note that another source of inflation expectations is the Consensus Forecasts monthly report. However, the expectations provided there are made in a fixed-horizon basis. This is, every month it is reported the forecast for December of the current and next year. Hence, the information provided for intermediate horizons would be weaker than that coming from a moving horizon forecast. More over, this will redound into an inefficient forecast since the implied errors will show smaller errors at longer horizons that those made at shorter horizons.

Table 1 displays some descriptive statistics of all the series, including the output gap which is described in the next subsection. Basically, its construction relies on the use of the Economic Activity Monthly Index (EAMI, index level 2013=100), which constitutes a monthly measure of GDP.[5] Note that the preferred transformation to achieve stationary in level series is the annual percentage change. This transformation is preferred because it is achieved stationarity according to the Augmented Dickey-Fuller test; it is an easy to interpret standard transformation; and matches the denomination of the ChSPF answers.

[4] Database freely available at http://www.bcentral.cl/eng/economic-statistics/series-indicators/index_ee.htm. See Pedersen (2010) for details.

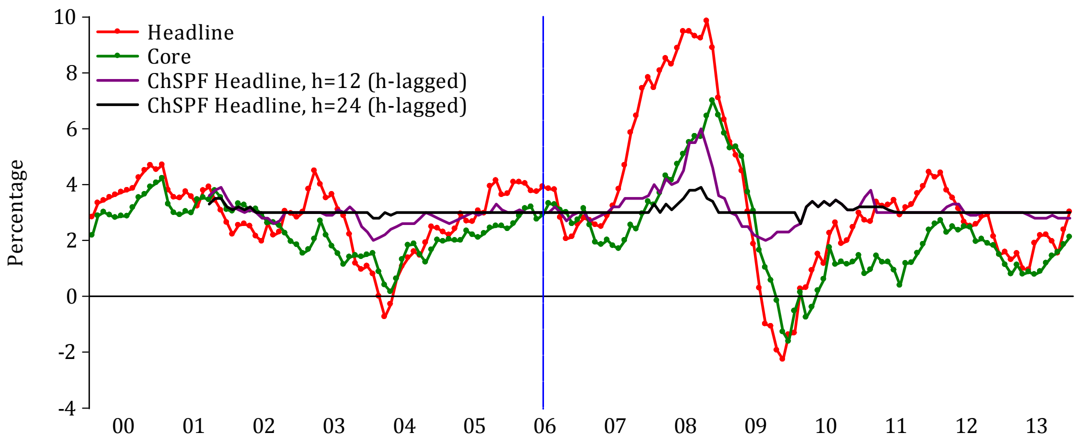
[5] Moreover, the annual rate of growth of the EAMI coincides with that of the GDP for each third month of each quarter. EAMI as well as inflation are freely available at: http://si3.bcentral.cl/Siete/secure/cuadros/arboles.aspx.

Table 1: Descriptive statistics of used time series (*)

	Symbol (*stationary*)	Mean	Median	Standard deviation	Max.	Min.	ADF Stat. (**) (*level*)	ADF Stat. (*annual var.*)
Inflation (Headline)	π_t	3.18	2.96	2.17	9.85	-2.27	-0.24 (0.930)	-2.59 (0.096)
Inflation (Core)	$\widetilde{\pi}_t$	2.32	2.22	1.42	7.00	-1.63	-2.94 (0.154)	-4.06 (0.009)
EAMI	y_t	4.40	4.67	2.63	13.18	-4.43	-2.80 (0.199)	-3.04 (0.033)
ChSPF: Inflation (*t+12*)	$\pi^f_{t,t+12}$	3.08	3.00	0.06	6.00	2.00	-3.99 (0.011)	-
ChSPF: Inflation (*t+24*)	$\pi^f_{t,t+24}$	3.07	3.00	0.17	3.90	2.60	-4.36 (0.003)	-
ChSPF: EAMI (*t+1*)	-	4.17	4.50	2.08	13.00	-3.60	-2.74 (0.069)	-
ChSPF: GDP (*T*) (***)	-	4.36	4.80	1.78	6.50	-1.80	-3.00 (0.037)	-
ChSPF: GDP (*T+1*)	-	4.80	5.00	0.46	6.00	3.30	-2.72 (0.074)	-
Output Gap Bwd.	$\widehat{y}_t$	-0.00	0.00	0.02	0.05	-0.06	-1.92 (0.053)	-
Output Gap Fwd. (*t+12*)	$\widehat{y}^f_{t,t+12}$	-0.00	-0.00	0.02	0.07	-0.07	-2.83 (0.005)	-
Output Gap Fwd. (*t+24*)	$\widehat{y}^f_{t,t+24}$	-0.04	-0.04	0.03	0.03	-0.09	-2.73 (0.072)	-

(*) Sample: 2000.1–2013.12 (168 obs.). (**) ADF stands for the Augmented Dickey-Fuller unit root test. ADF p-value shown in $(\cdot)$. ADF computed with constant, trend (Core, EAMI, ChSPF: Inflation (*t+12*), ChSPF: Inflation (*t+24*)), or none (Output Gap Bwd., Output Gap Fwd. (*t+12*)). Bandwidth ranging from 4 to 24 lags. (***) t stands for monthly frequency, while T for annual. Source: Author's elaboration.

Figure 1: Actual and h-lagged forecasted Headline and Core inflation (*)

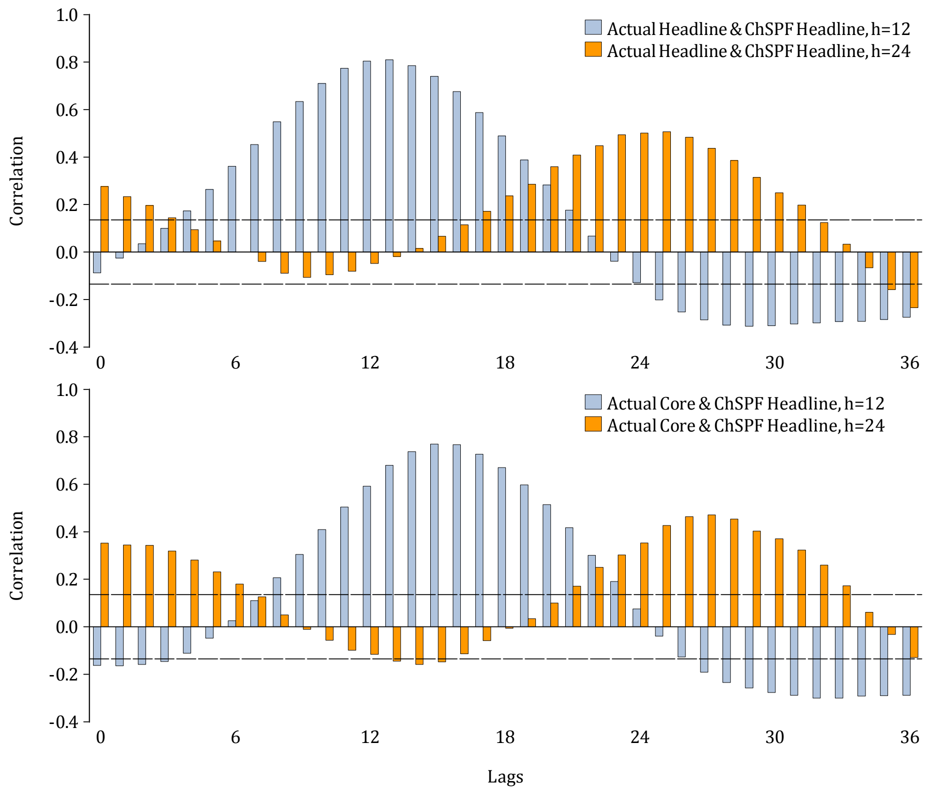


(*) Vertical line indicates out-of-sample forecasts start point (2006.6).
Source: Author's elaboration using CBC's dataset.

Figure 1 displays the actual and h-lagged forecasted inflation series across the whole sample. Note that the inflation expectation 24-months ahead ("ChSPF: Inflation (*t+24*)") is very close to the inflation target the majority of the time. Also, the time span includes the global inflationary spillover of the recent financial crisis.

Note that the use of ChSPF dataset is made under a number of implicit assumptions. One of the most important is that respondents minimise their mean squared forecasted error, *i.e.* quadratic loss function. This implies, among other results, that they are efficient into incorporating and using new available information. For an appraisal of the suitability of these projections, in Figure 2 I plot the cross-correlation between inflation (both) and the ChSPF expectations for 12 and 24 months. After noticing that the forecast is made for headline inflation, both expectations variables match the horizon at which they are targeting relatively well. As expected, however, it is a less clear cut with core inflation. In that case it is observed that expectations match the horizon with almost 3 or 4 lags but with a similar accuracy.

Figure 2: Cross-correlation. Inflation and (lags of) ChSPF expectations (*)



(*) Confidence interval: $0 \pm Z_\alpha/\sqrt{n}$, where α is the probability-level of the inverse normal distribution (n=168) (see Chatfield, 2004, for details). Source: Author's elaboration.

2.2 Output gap building blocks

One of the major drawbacks when estimating the NKPC is the impossibility to accurately measure the excess of demand–*i.e.* marginal costs. The typical alternative is the output gap–*i.e.* the difference between the current and potential output.[6] Basically, instability arise with the "end-of-sample" problem of filtering, especially when the Hodrick-Prescott (HP) procedure is used to obtain the potential output; an unobservable component.[7] To alleviate this setback, I follow the approach proposed by Kaiser and Maravall (1999). This consists of adding forecasted observations to level series prior to perform any filtering procedure. Hence, the method applied to obtain the output gap follows the steps of Figure 3. Note that the seasonal adjustment is made with X12-ARIMA in its default mode, and the filtering method is HP (λ=129,600).

As the method involves the use of forecasted observations, three measures of output gap emerges: (i) using forecasted values up to 5-years ahead (60 observations) coming from an ARMA(p, q) model (labelled "Bwd."), (ii) using ChSPF GDP forecast for the current year ("Fwd. $(t{+}12)$"), and (iii) same as (ii) but using forecast for the following year ("Fwd. $(t{+}24)$"). As a result, three different matched specifications of the model in (1) are analysed:

[6]Note that I focus on *output* gap instead of *unemployment* gap following the recommendations of Staiger, Stock, and Watson (1997a, 1997b).

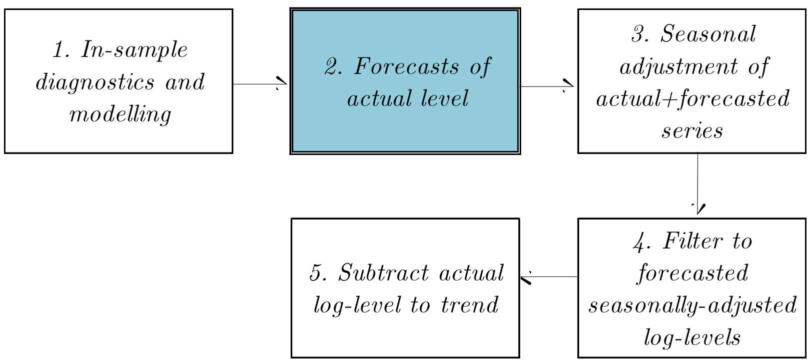
[7]See Orphanides (2001), Orphanides and van Norden (2002, 2005) and Garratt *et al.* (2008) for a discussion on this matter.

1. a (now *non-strictly*) BL model, including lagged inflation only, plus "Bwd." output gap,
2. a FL model, including lagged inflation, the ChSPF expectations of inflation 12-months ahead, plus "Fwd. (*t+12*)" output gap, and
3. a FL model, including lagged inflation, the ChSPF expectations of inflation 24-months ahead, plus "Fwd. (*t+24*)" output gap.

The chosen ARMA model for EAMI corresponds to $\Delta^{12}Y_t = y_t = \alpha + \rho y_{t-1} + \theta_1 \upsilon_{t-1} + \theta_{12}\upsilon_{t-12} + \upsilon_t$, with $\upsilon_t \sim iid\mathcal{N}(0, \sigma_\upsilon^2)$, chosen with the *General-to-Specific* (GETS) iterative process allowing for skipped terms. The estimation is presented in Table 2, which also reveals robust results across the sample span, and a correct specification according to the Durbin-Watson statistic.

In Appendix A it is compared the stability across the sample of the purely BL and "Bwd." output gap measures to assess the stability gain using forecast observations. This procedure redounds into a more demanding BL benchmark for the HNKPC estimation and forecasts. As expected, the latter methodology exhibit minor deviations while the number of observation is increased.

Figure 3: Output gap building blocks



Source: Author's elaboration.

Table 2: Auxiliary model for EAMI (y_t) forecasts (*)

	(1)	(2)
	Estimation sample	Full sample
Dep. variable	y_t	y_t
ρ	0.961	0.893
	(0.000)	(0.000)
θ_1	-0.510	-0.226
	(0.000)	(0.000)
θ_{12}	-0.489	-0.773
	(0.000)	(0.000)
α	6.536	4.360
	(0.000)	(0.000)
$\overline{R}^2$	0.656	0.741
D-W statistic	2.288	2.355
RMSE (**)	1.209	1.324
Sample	2000.2–2006.5	2000.2–2013.12
No. obs.	76	167

(*) p-value shown in $(\cdot)$. Variance corrected with Newey-West HAC. (**) RMSE stands for Root Mean Squared Error. Source: Author's elaboration.

2.3 Out-of-sample assessment

To investigate whether the BL or one of the two FL specifications is better at forecasting, I compute and compare the Root Mean Squared Forecast Error (RMSFE):

$$RMSFE_h = \left[\frac{1}{T} \sum_{t=1}^{T} (\pi_{t,t} - \pi^f_{t,t-h})^2 \right]^{\frac{1}{2}}, \tag{5}$$

where $\pi^f_{t,t-h}$ is the forecast h-step-ahead of $\pi_{t,t}$, made at period t. For completeness, and a more demanding comparison, I also include two competing models: the random walk (RWK), and an AR(p) model choosing p according to a fixed-T version of the *stepwise backwards* procedure (labelled "AR[SB]"). This last model, similar to GETS, chooses the autoregressive order p within the estimation sample, fixing it until the last observation is used for estimation. Note that OLS deliver misleading results (not shown), implying that each forecast involve the multistage estimation once an observation is added to the sample (and dropping the last one under a rolling window scheme).

Finally, statistical inference is carried out with the GW test of predictive ability. It requires that errors have to be computed in a rolling window scheme, and works for both nested and nonnested models. The null hypothesis can be summarised as *both models have the same predictive ability conditional to its model* (see Clark and McCracken, 2013, for a comprehensive description of the test.)

3 Results

3.1 In-sample results

The results for the three specifications with headline are presented in Table 3 for two samples: *estimation* (1–3) and *full sample* (4–6). The J-stat p-value indicates that IV are valid along the sample span except for the BL specification. The list of IV and its used lags is presented in Table 5. There are two other variables tested as IV: Consensus Forecasts' Brent oil price and ChSPF's foreign exchange rate. They both result as no valid IV with any acceptable lag length.

Note that in both BL equations ((1) and (4)), the lagged inflation coefficients ranged from 0.83 to 0.88 (both significant). The output gap is significant with one lag (note that the first lag is allowed as it comes from a forecasted variable. In reality, delay in data release allows since 2 lags onwards). Equation (2) is the preferred with "Fwd. *(t+12)*". In this case, the output gap is not significant with any lag between [1;24]. Equation (2') shows the results when considering the 12-lag. As the data for t are sorted considering the -h-period value, any lag between [1;12] can be still considered as a forecasted value of π_t (in this case, lag 12 matches the targeted variable). Nevertheless, the output gap results as a valid IV. The FL coefficient accounts from 1.08 times bigger than the lagged coefficients in the first sample (Equation 2), declining to 0.67 times with the whole sample (Equation 5). The set of equations (3), (3') and (6) mimics the results for "Fwd. *(t+24)*". In this case, the decay in importance of the FL coefficient is more dramatic. For the first sample (Equation 3) accounts for 1.58 times to then decay to 0.40 with the full sample (Equation 6).

Table 4 shows the results for core inflation. Qualitatively these results are similar to headline but quantitatively their figures are more dramatic. The lagged inflation coefficient in the BL specification fluctuates between 0.77 and 0.91 (Table 4: Equations 1 and 4). The FL coefficient in the "Fwd. *(t+12)*" specification starts from 2.48 times the lagged coefficient, declining to

0.39 when considering full sample. Considering the "Fwd. (*t+24*)", the FL coefficient accounts from 1.12 times with respect to the lagged, to just 0.19 with full sample.

All these results reveal instability in the parameters associated to FL inflation. To this end, in Figure 4 I display four graphs for each variable analysing the evolution across the sample (recursive) of the key parameters: γ_b, γ_f, the t-Statistic of γ_f, and the J-stat p-value (keeping the same IV).[8] These results show that for headline the persistence parameter moves slowly around 0.80 to 0.90 at the end of the sample. However, different results are obtained for the FL parameter. A major shift is adverted in the aftermath of the financial crisis. While in 2009 the parameter reaches values even greater than one, since 2012 that is around 0.50 with the two FL specifications. The parameter is almost always significant, and the IV are valid until 2013 for the FL specifications only.

For core inflation the situation looks similar. However, almost all estimates remain steady since late 2009. The lagged coefficients look similar for the three specifications around 0.90, while the FL coefficient below 0.40 (significant along the sample). The IV are consistent, especially with the "Fwd. (*t+24*)" specification.

From this analysis it is possible to conclude that there is a robust but low role for expectations when determining current inflation. This evidence is shared for headline as well as core inflation.

3.2 Out-of-sample results

The results are presented in terms of the "RMSFE ratio" between the preferred FL specification ("pivot") and a competing model:

$$RMSFE\ Ratio_h = \frac{RMSFE_h^{Fwd.\ (t+k)}}{RMSFE_h^{Competing}}.$$

Hence, figures below one are in favour of the "Fwd. (t+k)" model, where k=12 for headline and k=24 for core. The results are presented in Table 6.

The results for headline show predictive gains in almost all cases. The exceptions are with respect to the RWK and the AR[SB] at h={1;3}. Note that when comparing to the other PC, the gains are qualitatively mixed: while higher gains are observed respect to "Fwd. (*t+24*)" at h={1;3}, it achieves 45.9% (=1-0.541) when predicting at h={6;12}. The preferred specification is also better than both benchmarks when predicting at h={6;12}. According to the GW test, all differences are statistically significant except those with the BL specification.

The results for core reveals that the preferred specification "Fwd. (*t+24*)" outperforms the other FL specification, and both benchmarks when h=12. The GW test reveals that only respect to "Fwd (*t+12*)" at h={1;3} the gains are statistically significant. However, note the BL specification is better at any horizon (but gains not significant). This result suggests that the lower variance of core respect to headline–*i.e.* its smoothness–inflates the relevance of the autoregressive term neglecting the inflationary FL variable (recalling that the forecast is made for headline).

[8] However, this analysis is simpler than that developed, for instance, in Swamy and Tavlas (2007) and Hondroyiannis, Swamy, and Tavlas (2009). In those studies, the authors make use of a time-varying coefficient environment to reduce bias specification, finding a minor role for lagged inflation in four European countries.

Table 3: Estimation results for Headline inflation (*)

	(1)	(2)	(2')	(3)	(3')	(4)	(5)	(6)
Dep. variable:	*Headline inflation*: π_t							
	Estimation sample					Full sample		
π_{t-1}	0.829	0.750	0.802	0.772	0779	0.882	0.807	0.900
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\pi^f_{t,t+12}$	-	0.806	0.890	1.220	1.144	-	0.542	0.356
	-	(0.032) [12]	(0.008) [12]	(0.003) [9]	(0.004) [9]	-	(0.000) [12]	(0.069) [9]
$\widehat{y}_{t-1}$	0.210	-	-	-	-	0.135	-	-
	(0.004) [1]	-	-	-	-	(0.043) [1]	-	-
$\widehat{y}^f_{t,t+12}$	-	IV (**)	-0.290	-	-	-	IV	-
	-	-	(0.397) [12]	-	-	-	-	-
$\widehat{y}^f_{t,t+24}$	-	-	-	IV	-0.012	-	-	IV
	-	-	-	-	(0.712) [1]	-	-	-
Constant	0.543	-1.641	-2.200	-2.837	-2.702	0.400	-1.106	-0.699
	(0.001)	(0.075)	(0.016)	(0.008)	(0.007)	(0.000)	(0.004)	(0.180)
J-statistic	0.000	0.879	0.520	1.307	1.218	4.496	4.065	3.688
J-stat. p-value	(0.979)	(0.644)	(0.470)	(0.520)	(0.269)	(0.033)	(0.130)	(0.158)
Sample	2000.5–	2002.2–	2002.2–	2002.9–	2002.9–	2000.5–	2002.2–	2002.9–
	2006.5	2006.5	2006.5	2006.5	2006.5	2013.12	2013.12	2012.2
No. obs.	73	52	52	45	45	164	143	114

(*) p-value shown in $(\cdot)$; chosen lag shown in $[\cdot]$, both below the coefficient estimates. Estimations with GMM. Weighting matrix estimation: covariance matrix inverse (with Newey-West HAC). Whitening lag specification: automatic with BIC, allowing up to 3 lags. (**) IV stands for instrumental variable. Source: Author's elaboration.

Table 4: Estimation results for Core inflation (*)

	(1)	(2)	(2')	(3)	(3')	(4)	(5)	(6)
Dep. variable:	*Core inflation*: $\widetilde{\pi}_t$							
	Estimation sample					Full sample		
$\widetilde{\pi}_{t-1}$	0.768	0.526	0.650	0.645	0.885	0.914	0.867	0.939
	(0.000)	(0.031)	(0.033)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\pi^f_{t,t+12}$	-	1.303	1.034	0.725	0.361	-	0.336	0.175
	-	(0.106) [12]	(0.181) [12]	(0.034) [12]	(0.117) [1]	-	(0.000) [12]	(0.012) [12]
$\widehat{y}_{t-1}$	0.212	-	-	-	-	0.065	-	-
	(0.000) [1]	-	-	-	-	(0.030) [1]	-	-
$\widehat{y}^f_{t,t+12}$	-	IV	-0.082	-	-	-	IV	-
	-	-	(0.494) [2]	-	-	-	-	-
$\widehat{y}^f_{t,t+24}$	-	-	-	IV	-0.050	-	-	IV
	-	-	-	-	(0.048) [1]	-	-	-
Constant	0.634	-2.473	-2.302	-1.305	-1.090	0.217	-0.725	-0.351
	(0.005)	(0.146)	(0.166)	(0.073)	(0.038)	(0.008)	(0.000)	(0.051)
J-statistic	2.086	0.167	0.007	3.556	2.577	1.490	3.845	2.800
J-stat. p-value	(0.148)	(0.919)	(0.933)	(0.168)	(0.108)	(0.222)	(0.146)	(0.246)
Sample	2000.5–	2002.2–	2002.2–	2002.9–	2002.9–	2000.5–	2002.2–	2002.9–
	2006.5	2006.5	2006.5	2006.5	2006.5	2013.12	2013.12	2012.2
No. obs.	73	52	52	45	45	164	143	114

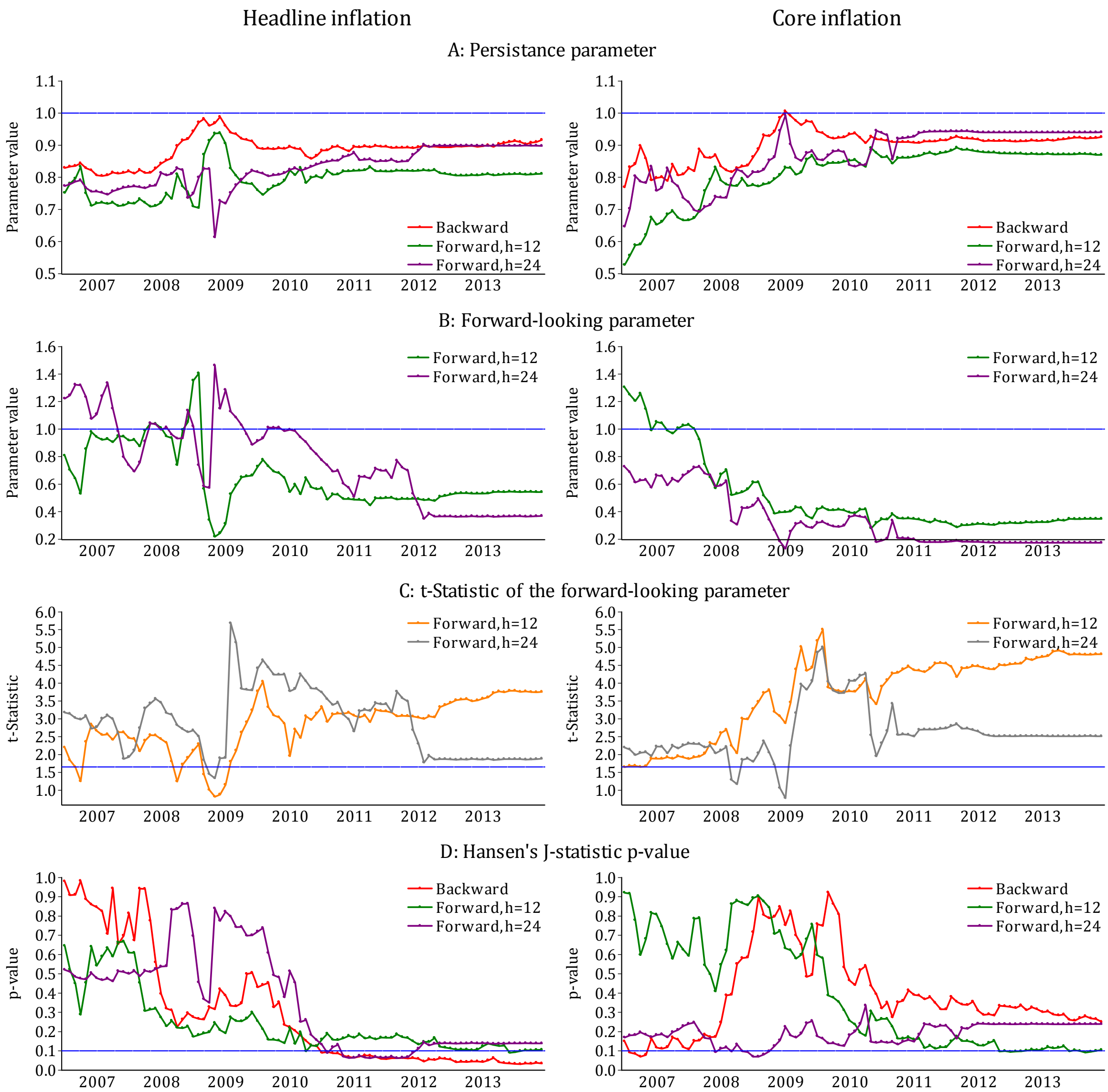
(*) See notes in Table 3. Source: Author's elaboration.

Table 5: Instrumental variables list

Equation	Instruments
Headline inflation, Table 3	
(1), (4)	$Constant,\ \pi_{t-3},\ \pi_{t-4},\ \widehat{y}_{t-3}$
(2), (2'), (5)	$Constant,\ \pi_{t-3},\ \pi^f_{t-24,t+24},\ \widehat{y}^f_{t-12,t+12},\ \widehat{y}^f_{t-25,t+12}$
(3), (6)	$Constant,\ \pi_{t-3},\ \pi^f_{t-24,t+24},\ \widehat{y}^f_{t-2,t+24},\ \widehat{y}^f_{t-20,t+24}$
Core inflation, Table 4	
(1), (4)	$Constant,\ \widetilde{\pi}_{t-3},\ \widetilde{\pi}_{t-4},\ \widehat{y}_{t-2}$
(2), (2'), (5)	$Constant,\ \widetilde{\pi}_{t-3},\ \pi^f_{t-24,t+24},\ \widehat{y}^f_{t-12,t+12},\ \widehat{y}^f_{t-25,t+12}$
(3), (6)	$Constant,\ \widetilde{\pi}_{t-3},\ \pi^f_{t-24,t+24},\ \widehat{y}^f_{t-2,t+24},\ \widehat{y}^f_{t-20,t+24}$

Source: Author's elaboration.

Figure 4: In-sample results of recursive parameter estimation across forecasting sample (*)



(*) A and B: Horizontal line=unit root bound. C: Horizontal line=Z_α^{-1}, where α is the probability-level (10%) of the inverse normal distribution. D: Horizontal line: p-value=10%. Source: Author's elaboration.

Table 6: Out-of-sample results. RMSFE ratio (*)

	Headline inflation					*Core inflation*					
	(1)	(2)	(3)	RWK	AR[SB] (*)	(1)	(2)	(3)	RWK	AR[SB]	No. obs.
h=1	0.966	1.000	0.791*	7.757	9.360	2.507	0.707**	1.000	10.300	10.865	91
h=3	0.716	1.000	0.636***	1.242	1.511	2.162	0.721**	1.000	2.454	2.576	89
h=6	0.507	1.000	0.605***	0.373**	0.416**	1.901	0.815	1.000	0.980	1.099	86
h=12	0.541	1.000	0.787**	0.177**	0.193**	2.359	0.909	1.000	0.534	0.595	80

(*) RMSPE ratio stands for RMSPE(*Pivot*)/RMSPE(*Competing*). GW test results: (***) $p<1\%$, (**) $p<5\%$, (*) $p<10\%$. Figures below 1 in yellow; pivot in grey. (**) AR[SB] stands for *stepwise backward* model selection; 3 lags chosen for Headline and Core inflation. Source: Author's elaboration.

In general, the out-of-sample exercise suggests that along with the ability of the HNKPC to explain inflation dynamics, it could be also considered as a valid benchmark model when forecasting at short-run. The predictive results for core inflation point out that its dynamics differs from those of headline, suggesting that core could be a process with higher memory (Granger and Joyeux, 1980). It is also suggested that the FL measures used are more related to the most volatile components of inflation. Conditional to the IV, the output gap measure plays a role within the BL specification delivering better results than its closer benchmark, AR[SB]. Finally, unexplored vignettes in this article may shed some light on core dynamics by analysing some minor twists. For instance, nonlinearities in the (same) IV, and/or long-run forecasting horizons.

4 Concluding remarks

The aim of this article is to investigate to which extent FL measures of inflation help to explain inflation dynamics and their forecasts with a PC ensemble. This objective is tackled by analysing the performance of the HNKPC, using a dataset of the Chilean economy, including inflation forecasts as a measure of inflation expectations.

To that end, I first estimate with GMM an unrestricted version of the HNKPC, to then compare its predictive power with a BL PC and traditional benchmarks predicting at $h = \{1; 3; 6; 12\}$-months-ahead.

The results show that the FL inflationary component is statistically significant when is included in the specification. In size, the preferred specification accounts from 1.58 to 0.40 times the lagged inflation coefficient; the latter figure considering whole sample. When considering short-term forecasting, I find predictive gains close to 45% (respect to the BL specification) and up to 80% (respect to the RWK) when forecasting at 12-months-ahead. However, these gains are not statistically significant. In sum, these results should be read carefully and the HNKPC just as a valid benchmark.

Finally, I make use of core inflation measure to conduct a robustness analysis. In-sample results support the existence of the HNKPC. Nevertheless, predictive results suggest that core could be a process with higher memory. However, the output gap plays a key role delivering better results than similar benchmark. It is also suggested from the results for core inflation that FL measures are related to most volatile components of inflation. Nevertheless, more research on those linkages is left for further research.

Acknowledgements

I thank the comments and suggestions to Rolando Campusano, Tim Lloyd, Pablo Medel, and Damián Romero. Nevertheless, I exclude them for any error or omission that remains at my own responsibility.

References

1. Brissimis, S.N. and N.S. Magginas, 2008, "Inflation Forecasts and the New Keynesian Phillips Curve," *International Journal of Central Banking* **08**(June): 1-22.

2. Calvo, G.A., 1983, "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics* **12**(3): 383-98.

3. Céspedes, L.F., M. Ochoa, and C. Soto, 2005, "*The New Keynesian Phillips Curve in an Emerging Market Economy: The Case of Chile*," Working Paper 355, Central Bank of Chile.

4. Chatfield, C., 2004, *The Analysis of Time Series: An Introduction*, Sixth Edition, Chapman and Hall/CRC Texts in Statistical Science.

5. Christiano, L.J., M. Eichenbaum, and C.L. Evans, 2005, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy* **113**(1): 1-45.

6. Clark, T. and M. McCracken, 2013, *Advances in Forecast Evaluation*, in A. Timmermann and G. Elliot (eds.), *Handbook of Economic Forecasting, Volume 2*, Elsevier, North-Holland.

7. Collard, F. and H. Dellas, 2004, "*The New Keynesian Model with Imperfect Information and Learning*," Working Paper 273, Institut d'Économie Industrielle (IDEI), Toulouse, France.

8. Croushore, D. and T. Stark, 2001, "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics* **105**(1): 111-30.

9. Erceg, C.J. and A.T. Levin, 2003, "Imperfect Credibility and Inflation Persistence," *Journal of Monetary Economics* **50**(4): 915-44.

10. Findley, D.F, B.C. Monsell, W.R. Bell, M.C. Otto and B.-C. Chen, 1998, "New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program," *Journal of Business and Economic Statistics* **16**(2): 127-152.

11. Findley, D.F., B.C. Monsell, H.B. Shulman and M.G. Pugh, 1990, "Sliding Spans Diagnostics for Seasonal and Related Adjustments," *Journal of the American Statistical Association* **85**(410): 345-355.

12. Galí, J. and M. Gertler, 1999, "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics* **44**(2): 195-222.

13. Galí, J., M. Gertler, and J.D. López-Salido, 2005, "Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve," *Journal of Monetary Economics* **52**(6): 1107-18.

14. Garratt, A., K. Lee, E. Mise, and K. Shields, 2008, "Real-Time Representations of the Output Gap," *The Review of Economics and Statistics* **90**(4): 792-804.

15. Giacomini, R. and H. White, 2006, "Tests of Conditional Predictive Ability," *Econometrica* **74**(6): 1545-78.

16. Granger, C.W.J. and R. Joyeux, 1980, "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis* **1**: 15-29.

17. Hansen, L.P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* **50**(4): 1029-54.

18. Hondroyiannis, G., P.A.V.B. Swamy, and G.S. Tavlas, "The New Keynesian Phillips Curve in a Time-Varying Coefficient Environment: Some European Evidence," *Macroeconomics Dynamics* **13**: 149-166.

19. Jean-Baptiste, F., 2012, "Forecasting with the New Keynesian Phillips Curve: Evidence from Survey Data," *Economics Letters* **117**(3): 811-13.

20. Kaiser, R. and A. Maravall, 1999, "Estimation of the Business Cycle: A Modified Hodrick-Prescott Filter," *Spanish Economic Review* **1**: 175-206.

21. Orphanides, A., 2001, "Monetary Policy Rules Based on Real-Time Data," *American Economic Review* **91**(4): 964-85.

22. Orphanides, A. and S. van Norden, 2002, "The Unreliability of Output-Gap Estimates in Real Time," *The Review of Economics and Statistics* **LXXXIV**(4): 569-83.

23. Orphanides, A. and S. van Norden, 2005, "The Reliability of Inflation Forecasts based on Output Gap Estimates in Real Time," *Journal of Money, Credit, and Banking* **37**(3): 583-601.

24. Paloviita, M. and D. Mayes, 2005, "The Use of Real-Time Information in Phillips-Curve Relationships for the Euro Area," *The North American Journal of Economics and Finance* **16**(3): 415-34.

25. Pedersen, M., 2010, "*An Introductory Note to the Survey of Professional Forecasters*," [in Spanish] Studies in Economic Statistics 82, Central Bank of Chile.

26. Pincheira, P. and H. Rubio, 2010, "*The Low Predictive Power of Simple Phillips Curves in Chile: A Real-Time Evaluation*," [in Spanish] Working Paper 559, Central Bank of Chile.

27. Roberts, J.M., 1997, "Is Inflation Sticky?," *Journal of Money, Credit, and Banking* **39**(2): 173-96.

28. Rudd, J. and K. Whelan, 2005, "New Tests of the New-Keynesian Phillips Curve," *Journal of Monetary Economics* **52**(6): 1167-81.

29. Smets, F. and R. Wouters, 2003, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* **1**(5): 1123-75.

30. Smets, F. and R. Wouters, 2005, "Comparing Shocks and Frictions in US and Euro Area Business Cycles: A Bayesian DSGE Approach," *Journal of Applied Econometrics* **20**(2): 161-83.

31. Staiger, D., J.H. Stock, and M.W. Watson, 1997a, *How Precise are Estimates of the Natural Rate of Unemployment?*, in C. Romer and D. Romer (eds.), *Reducing Inflation: Motivation and Strategy*, Chicago University Press.

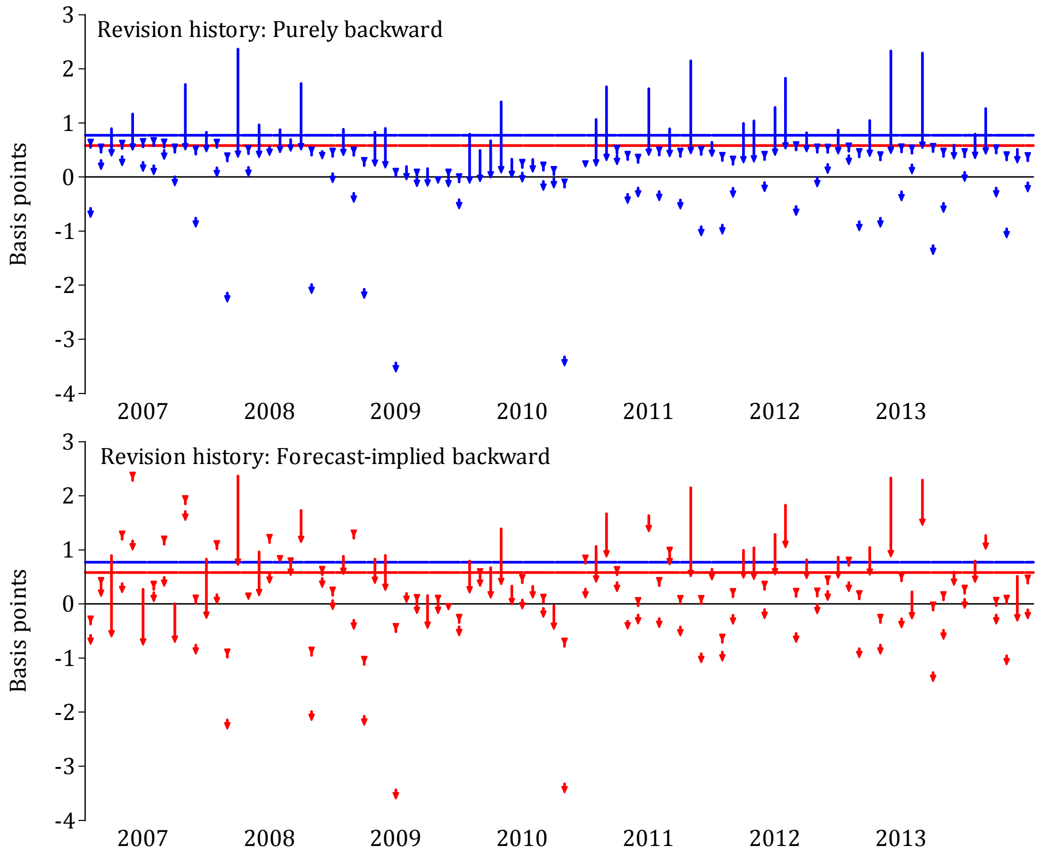
32. Staiger, D., J.H. Stock, and M.W. Watson, 1997b, "The NAIRU, Unemployment and Monetary Policy," *Journal of Economic Perspectives* **11**(1): 33-49.

33. Swamy, P.A.V.B. and G.S. Tavlas, 2007, "The New Keynesian Phillips Curve and the Inflation Expectations: Re-specification and Interpretation," *Economic Theory* **31**: 293-306.

A Output gap stability analysis

One of the most desirable conditions for an unobservable variable is its stability. This can be understand as how robust is the measure while more observations are added to the sample. A more robust measure is that less invariant to new observations, and statistical inference can be carried out with a higher degree of reliability.

Figure A1: Revision history comparison



▼=Most recent. Source: Author's elaboration.

There are several measures towards stability assessment. Some common as well as useful measures are those contained in the X12-ARIMA program in order to assess the seasonal adjustment quality, *i.e.* *sliding spans* and *revision history*.[9] In this appendix it is described and employed the revision history technique to determine the effect of forecast observations in the stability of the output gap measure, compared with the case where no observations are added. This last situation is often referred as the "end-of-sample" identification problem.

The revision history is defined as the difference between the earliest estimation of a given observation obtained when that observation is the last available and a later estimation based on all future data available at the time. Hence, this measure is specifically concerned with the effect of new information on the historical record of the output gap and the variance contribution to the estimation and the forecast afterwards.

The revision history is calculated as follows. Let $\widehat{y}_{t|t} = y_{t|t} - y^{\tau}_{t|t}$ the output gap measure (in logs) calculated using $y^{\tau}_{t|t}$ as a measure of potential output. $y^{\tau}_{t|t}$ corresponds to the trend component

[9] See Findley *et al.* (1990) and Findley *et al.* (1998) for details.

of the decomposition $y_{t|t} = y^{\tau}_{t|t} + y^{c}_{t|t}$, obtained with the HP filter using available data until observation t. Now, suppose that the same $\widehat{y}_{t|t}$ measure is obtained considering all future data available until observation T, $\widehat{y}_{t|T}$. The revision history is defined as:

$$R_t = \widehat{y}_{t|T} - \widehat{y}_{t|t}. \tag{A1}$$

Note also that the decomposition $y_{t|t} = y^{\tau}_{t|t} + y^{c}_{t|t}$ can be made by using the actual plus h-forecast-augmented variable, $y^{f}_{t|t+h}$, to improve its stability. In this case, the output gap corresponds to $\widehat{y}_{t|t,f} = y_{t|t} - y^{f,\tau}_{t|t+h}$, while the revision history to:

$$R_{t,f} = \widehat{y}_{t|T} - \widehat{y}_{t|t,f}. \tag{A2}$$

The comparison comprises R_t and $R_{t,f}$, as R_t is related to the purely BL case and $R_{t,f}$ to the "Bwd." output gap measure. In Figure 1A, the first panel show the revision history across the sample for output gap based on the purely BL potential output ($\blacktriangledown$-point is the "most recent" estimation $\widehat{y}_{t|T}$). The second panel exhibit the revision history for "Bwd.". In both figures there is also depicted the average of both measures. Note that the difference between purely BL and "Bwd." accounts for approximately 0.20 ($\simeq$0.78–0.59) basis points, while the variances are 0.83% and 0.59%, respectively. Hence, the procedure proposed by Kaiser and Maravall (1999) of adding forecast observations prior to any filtering procedure deliver a more stable measure of output gap. This last characteristic is desirable since this variable is prone to exhibit a larger measurement error which may turns to spoiling both interpretation and inference.