Forecasting Chilean Inflation with the Hybrid New Keynesian Phillips Curve: Globalisation, Combination, and Accuracy

Medel, Carlos A.

Central Bank of Chile

16 April 2017

Forecasting Chilean Inflation with the Hybrid New Keynesian Phillips Curve: Globalisation, Combination, and Accuracy*

Carlos A. Medel†
Financial Stability Area
Central Bank of Chile

April 16, 2017

Abstract

This article analyses the multihorizon predictive power of the *Hybrid New Keynesian Phillips Curve* (HNKPC) covering the period from 2000.1 to 2014.12, for the Chilean economy. A distinctive feature of this article is the use of a Global Vector Autoregression (GVAR) specification of the HNKPC to enforce an open economy version. Another feature is the use of direct measures of inflation expectations–*Consensus Forecasts*–differing from a fully-founded rational expectations model. The HNKPC point forecasts are evaluated using the Mean Squared Forecast Error (MSFE) statistic and statistically compared with several benchmarks, including combined forecasts. The results indicate that there is evidence to do not reject the hypothesis of the HNKPC for the Chilean economy, and it is also robust to alternative specifications. In predictive terms, the results show that in a sample previous to the global financial crisis, the evidence is mixed between atheoretical benchmarks and the HNKPC by itself or participating in a combined prediction. However, when the evaluation sample is extended to include a more volatile inflation period, the results suggest that the HNKPC (and combined with the random walk) delivers the most accurate forecasts at horizons comprised within a year. In the long-run the HNKPC deliver accurate results, but not enough to outperform the candidate statistical models.

JEL-Codes: *C22*; *C26*; *C53*; *E31*; *E37*; *E47*.

Keywords: *New Keynesian Phillips Curve*; *inflation forecasts*; *out-of-sample comparisons*; *survey data*; *Global VAR*; *structured time-series models*; *forecast combinations*.

*I thank the early-stage insights and comments of Rodrigo Caputo (Editor), Carla Fucito, Kevin C. Lee, Pablo Medel, James H. Stock, two anonymous referees, seminar participants at the Central Bank of Chile, and Consuelo Edwards for editing services. I also thank the *International Markets Surveillance Group* at the *Financial Stability Area* of the Central Bank of Chile the opportunity to develop and support this research project. I exclude all of them for any error or omission. This article is dedicated to the memory of the late Victor A. Fucito. The views and ideas expressed in this paper do not necessarily represent those of the Central Bank of Chile or its authorities.

†E-mail: cmedel@bcentral.cl.

1 Introduction

The *Hybrid New Keynesian Phillips Curve* (HNKPC) consists in a relationship between inflation and economic slack, considering at the same time lagged and expected values of inflation. In this article, the multihorizon predictive power of the HNKPC in two versions for the Chilean inflation is analysed. These versions are a *closed-economy* specification plus a trade-partners-augmented version, becoming an *open-economy* specification. To the latter, a Global Vector Autoregression (GVAR) ensemble (Pesaran, Schuermann, and Weiner, 2004) is used.[1] The latter consist in the first attempt which explicitly consider a large number of trade partners in the same econometric specification to forecast Chilean inflation. These economics-based forecasts are compared with traditional time-series benchmarks used in the literature, plus three combined forecasts. Another distinctive element of this article is the use of direct measures of inflation expectations embedded in the two versions of the HNKPC for the Chilean inflation forecasting purposes.

The Chilean case is chosen as it represents a small open economy under an inflation targeting regime with a floating exchange rate and permeable to specific shocks. For instance, being located in South America–and the only country within the *Organisation for Economic Co-operation and Development* of the region–, it is subject to regional shocks originated mainly in the biggest regional economy (Brazil), and with strong trade and financial connections with the US, the Euro Zone, and China as the major trade partners. Hence, the GVAR includes its main trade partners making up 70% of its total trade.[2]

The analysed monthly sample covers from 2000.1 to 2014.12 (180 observations), comprising an *evaluation* sample of 108 observations from 2006.1 to 2014.12. The results indicate that there is evidence to do not reject the hypothesis of the HNKPC for the Chilean economy, *i.e.* that the lagged and expected inflation coefficients are statistically significant, as is also that of the output gap. This finding is obtained with the closed-economy version and robust to alternative specifications of the output gap.

The GVAR open-economy version also complies with the required statistical and economics-based tests. This implies that trade partners already help to explain domestic inflation in a richer econometric setup that allows for many simulation and scenario analysis. Moreover, practical implication of comparing closed- versus -open-economy results reveals the predictive gain when including richer foreign information with ease. In predictive terms, the results show that previous to the financial crisis the evidence is mixed between atheoretical models and the HNKPC by itself, or in a combined prediction. However, when the evaluation sample is extended to a more volatile period, the results suggest that the HNKPC in its two versions (and combined with the random walk model, RW) delivers the most accurate forecasts at horizons comprising a year. In the long run the combination between the closed-economy HNKPC and the RW delivers more accurate results than the benchmark, although not enough to outperform any purely statistical prediction. It is hence obtained that at short horizons, and when inflation increases its volatility, the HNKPC result in the best forecasting option compared to traditional statistical models; a finding that is reverted at longer horizons.

It is widely recognised that accurate forecasts are *ex ante* a key element for the success of almost all macroeconomic policies. For the case of policymakers concerning price stability under an inflation targeting regime, timely accurate inflation forecasts are crucial for the success of monetary policy. As expected, from time to time new inflationary risks emerge challenging both policymakers and the current

[1] In particular, it corresponds to an extension of the forecasting exercise described in Medel (2015a), applied to the case analysed in Medel (2015b).

[2] The remaining considered countries are Brazil (BRA), China (CHI), the Euro Zone (EUR), Japan (JPN), and the US (US).

methodological tools developed to understand inflation dynamics. These challenges threat inflation from a different point of view rather than those economies in a more dominant position, and especially for policymakers of small open economies, like Chile. In particular, imported inflation from commodities and trade partners plus the contagion of shocks from industrialised countries are of special interest. This consists in the main motivation for the inclusion of an open economy version of the HNKPC.

For the particular case of monetary policy, the challenge of modelling external inflationary pressures has to deal also with the link between past and future domestic inflation rates. This link reflects the traditional inertia exhibited by backward-looking price setter firms and a forward-looking component provided by rational expectations agents' behaviour. One successful proposal in this regard is the HNKPC, introduced by Galí and Gertler (1999), and analysed further in Galí, Gertler, and López-Salido (2001, 2005). Note that their proposal results in a convenient specification for forecasting purposes and allowing many price settings.[3]

The rest of the article proceeds as follows. Section 2 reviews the relevant literature concerning the many topics that converge in this article. These are statistical versus economics-based inflation forecasts with uni- and multi-variate models, particularly for small open economies. Section 3 provides a full description of the econometrics methods used for the HNKPC-based forecasts. It is also defined the in-sample strategy to determine which will be the specifications used for prediction. Also, a detail of the statistical inference carried out is provided for the out-of-sample results plus some robustness exercises. Section 4 presents the results divided into estimation diagnostics, robustness results, and forecast accuracy. Section 5 concludes.

2 Literature review

The quest for accurate inflation forecasts has a long tradition in macroeconometrics and central banking literature. Given that inflation typically presents a close-to-unity behaviour, its modelling has concerned many econometric issues with economic implications. There are two broad views of forecasting inflation: the atheoretical statistical manner, and the economics-based procedure.[4] However, the literature concerning emerging countries, and particularly the Chilean economy, is disproportionally less than that devoted to industrialised economies.

The atheoretical or statistical manner refers to the case where the prediction comes from a model without economic fundamentals, and the appropriate model is obtained purely based on statistical tests' results. In this article, I used a stationary autoregression (AR) and the single exponential smoothing (ES) forecasts. Some references on the use of these models can be found in Medel (2015a) and the references therein.

When inflation is forecast with economic models, the task is typically performed with a Phillips Curve specification. Yet far from the original model the basic foundation still remains. This is a trade off between an activity measure and a price level.[5] The HNKPC, however, includes more economic elements since it is derived from an optimisation problem in the style of modern macroeconomics. It was introduced by Galí and Gertler (1999) and extended in Galí, Gertler, and López-Salido (2001, 2005). Closer literature analysing the HNKPC can be found in Sbordone (2002), Smets and Wouters (2003, 2007), Levin *et al.*

[3] Some theoretical derivations of the HNKPC can be found in Smets and Wouters (2003, 2005), Christiano, Eichenbaum, and Evans (2005), Erceg and Levin (2003), and Collard and Dellas (2004), among others.

[4] A recent survey of the many inflation forecasting methods can be found in Faust and Wright (2014).

[5] An interesting exercise is conducted in Granger and Jeon (2011) where it is studied how the original Phillips Curve paper could be estimated with the time-series econometrics known 50 years later. This is made using the same original variables and sample, and providing some extensions for robustness.

(2005), and Rabanal and Rubio (2005). Some articles using direct measures of expectations are Paloviita and Mayes (2005) using *Consensus Forecasts* for 11 European countries, Nason and Smith (2008) for the US–using the *Survey of Professional Forecasters* (SPF)–, Henzel and Wollmershauser (2008)–using *CESifo World Economic Survey* for Italy–, Paloviita (2009) for the Euro Zone, and Medel (2015b) for Chile–using the Central Bank of Chile's SPF.

The majority of the HNKPC estimations concern developed countries and in different versions; see Medel (2015a) for a review. For the case of Chile, little research has been conducted on this matter. Some exceptions are Céspedes, Ochoa, and Soto (2005) and Pincheira and Rubio (2015). The first article derives a NKPC from a structural microfounded model, and analyses their in-sample ability to explain inflation dynamics. The second article addresses the issue of the weak predictive power of a purely backward-looking Phillips Curve with real-time data. While Céspedes, Ochoa, and Soto (2005) also provide an out-of-sample assessment, it is not the major concern of the authors. Instead, inner motivation of Pincheira and Rubio (2015)–shaping the specification search exercise–is precisely forecast accuracy.

In a recent study, Medel (2015b) analyses the case of forecast Chilean inflation with a single country HNKPC specification using the Central Bank of Chile's SPF. It is worth mentioning that despite that the single-country HNKPC predicts better than the alternatives, the evidence is weak on the existence of a Phillips Curve when using *core inflation*; hence, an alternative not explored in this article. Moreover, when the same output gap measure used in this article is replaced by the annual growth of an economic activity index that mimics GDP in a monthly frequency, the results are still in favour of the proposed forecast-implied output gap variable.

The open economy version of the HNKPC used in this article is built in a GVAR ensemble. Obviously, the GVAR is not the first attempt to explicitly link world areas and countries, but it keeps the number of estimated coefficients to a minimum, avoiding the *curse of dimensionality* traditionally associated to VAR estimations. The potential applications of the GVAR methodology by far outreaches the exercises found in the literature. The introduction of the GVAR by Pesaran, Schuermann, and Weiner (2004) provides an application estimating the effect of economic shocks on firms' conditional loss distributions using 25 countries grouped into 11 regions. In this article, however, it is used a compact-scale version of the main Chilean trade partners to evaluate the capacity of the GVAR to transform foreign information into forecast accuracy. For this task, there is no need to include the full range of available economies.

The exercise analysed in this article compares the predictive ability of the HNKPC in a single-country and a GVAR version. An intermediate result is to compare both specifications between them to provide robustness to a particular finding of Medel (2015b). This consists in the use of trade-related variables in the closed-economy version of the HNKPC which come out as non-statistically significant. Hence, no role was found for openness or trade variables. This article, which makes use of a different inflation expectation measure, analyses the role of the RER, also finding it as a non-significant variable. Consequently, the use of the GVAR in this article results in a new attempt in search for a role of openness in forecasting accuracy. However, a comparison between close- and open-economy versions of the HNKPC should be carefully analysed, since an open-economy version typically redounds in the inclusion of more variables in the model.

3 Econometric approach

In this section all forecasting models are described: single-country HNKPC (closed economy; CE-HNKPC) and the GVAR HNKPC (open economy, OE-HNKPC). The atheoretical models AR, RW, and ES are described in Appendix A following the same presentation given in Medel (2015a, 2015b). As part of the methodological procedures used for out-of-sample statistical inference, the Root Mean Squared Forecast Error Ratio (RMSFE Ratio) is defined as well as the Giacomini and White (2006; GW) testing procedure. In Appendix B and C the dataset and the output gap building blocks are fully described, respectively.

3.1 Closed economy: single-country HNKPC

To sketch its foundations, assume a staggered price-setting scheme *á la* Calvo (Calvo, 1983). Let $1-\theta$ be the fraction of firms that change prices in a given period, and $1-\omega$ the fraction of firms that set prices optimally in a forward-looking manner. Hence, the HNKPC consists of a weighted average between past and future values of inflation plus a driving process $\widetilde{y}_t$, leading to the HNKPC baseline equation:

$$\pi_t = \gamma\widetilde{y}_t + \lambda_b\pi_{t-1} + \lambda_f\mathbb{E}_t[\pi^f_{t,t+h}] + \varepsilon_t, \tag{1}$$

where π_t is inflation, $\mathbb{E}_t[\pi^f_{t,t+h}] = \widetilde{\pi}_t$ is the inflation expectation at period f measured with a forecast made h-steps-ahead at period t, and $\widetilde{y}_t$ is a real marginal cost measure. $\{\gamma, \lambda_b, \lambda_f, \sigma^2_\varepsilon\}$ 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 coming from the optimisation problem of a structural NKPC where:

$$\begin{aligned} \lambda_b &= \frac{\omega}{\phi}, \\ \lambda_f &= \frac{\beta\theta}{\phi}, \\ \gamma &= \frac{[(1-\omega)(1-\theta)(1-\beta\theta)]}{\phi}, \\ \phi &= \theta + \omega\left[1-\theta(1-\beta)\right], \end{aligned} \tag{2}$$

and β is a discount factor. To avoid part of the simultaneity in the variables of the right-hand side, equation (1) is estimated with the Generalised Method of Moments (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 same variables as instrumental variables (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 $\mathbf{m}_t$, containing ℓ IV ($\ell \geq k$) which fulfils:

$$\mathbb{E}_{t-1}[(\pi_t - \gamma\widetilde{y}_t - \lambda_b\pi_{t-1} - \lambda_f\mathbb{E}_t[\pi^f_{t,t+h}]]) \times \mathbf{m}_{t-1}] = 0. \tag{3}$$

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

$$J(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{w}}_T) = \frac{1}{T}(\pi_t - \mathbf{x}'_t\widehat{\boldsymbol{\beta}})'_t\mathbf{m}\widehat{\mathbf{w}}_T^{-1}\mathbf{m}'(\pi_t - \mathbf{x}'_t\widehat{\boldsymbol{\beta}}), \tag{4}$$

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

$$\widehat{\boldsymbol{\beta}} = (\mathbf{x}'\mathbf{m}\widehat{\mathbf{w}}_T^{-1}\mathbf{m}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{m}\widehat{\mathbf{w}}_T^{-1}\mathbf{m}'\pi_t, \tag{5}$$

that minimises equation (4). As $J(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{w}}_T) \sim \chi^2_{\ell-k}$, along with the estimated coefficients I also report 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 $\alpha\%$-level of significance, and the specification qualifies to be the forecasting model.

The estimation of the weighting matrix is made according to the 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 and West (1987) heteroskedasticity and autocorrelation correction (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 to the Bayesian Information Criterion (BIC) choosing in a maximum of 3 lags (following the $T^{1/3}$ rule).

All the estimations are made through the GMM estimator to find a particular specification using the estimation sample, and following a *General-to-Specific* (GETS) strategy for the first stage regression. There are many reasons to prefer GMM as the estimation method. First, and following Galí, Gertler, and López-Salido (2005), the GMM results are robust to the Non Linear IV GMM (NLIVGMM) estimator, which has been criticised by, for instance, Lindé (2005) and Rudd and Whelan (2005). This is a good reason to keep GMM since NLIVGMM estimation requires more computer time and it is more sensitive to the IV election in an univariate ensemble. However, to perform the forecasting estimations, I use the Ordinary Least Squares (OLS) estimator following the same methodology used by Jean-Baptiste (2012) for the UK, and Medel (2015a) for six major industrialised economies.[6] As emphasised by Cochrane (2001), the choice between one (GMM) or another maximum likelihood estimator for univariate cases is a trade-off, and no consensus has been achieved.

3.2 Open economy: Global VAR HNKPC

The use of the GVAR obeys particularly to an open economy version of the HNKPC. Galí and Monacelli (2005) develop an open economy version of the HNKPC which explicitly includes the interaction of a *domestic* country with the rest of the world. This is made through the real exchange rate (RER) and certain commodity prices in the output gap measure. The model is based on a richer economic environment but delivering a reduced-form specification including domestic inflation and output gap also suitable for forecasting exercises. Nevertheless, there is neither a unique nor a widely accepted manner in which a foreign component may be considered in the HNKPC. The option provided by the GVAR is to include an international trade-partners-related version of the same variables used to model the close economy case. Hence, the GVAR naturally extends any close economy estimation into another in which all the countries (or regions) are interconnected with one another.[7]

The GVAR methodology was introduced by Pesaran, Schuermann, and Weiner (2004) in search for a flexible procedure able to include key interactions across a big number of countries. Model flexibility comes from the fact that it is possible to model a country-level VAR including specific variables and different lag length. The foreign variables enter in the domestic equation as weighted averages of the same variables defined for the remaining countries. As the weights are exogenously imposed it is easy to define first the model in a "compressed" manner, making possible its estimation, to then "decompress" it for further post-estimation handling. Given the mechanics of the GVAR, it avoids the *curse of dimensionality* confronted by VAR models with too many coefficients to be estimated (and exponentially arisen when a new variable is included).

For formal description purposes (following closely Pesaran, Schuermann, and Weiner, 2004), assume that there are i=0,1,...,N+1 countries across the time t=1,...,T, where the country i=0 is the reference country. Now, assume that each country is modelled using k_i domestic and k_i^* foreign variables (hereafter, "*"

[6] Empirical results do not deliver substantial parameter differences between GMM and OLS.

[7] I also analyse the role of the RER dynamics into the single-country HNKPC, which can be understood as an intermediate step between the baseline HNKPC and the GVAR specification.

will refer to foreign variables). In this article, for each country $k_i{=}k_i^*{=}3$, and hence $k{=}6$ (accounting: $k_i{=}\{\pi_{i,t-1}, \widetilde{\pi}_{it}, \widetilde{y}_{it}\}$ and $k_i^*{=}\{\pi^*_{i,t-1}, \widetilde{\pi}^*_{it}, \widetilde{y}^*_{it}\}$). So, for each country i it is defined the $k_i \times 1$ vector $\mathbf{x}_{it} = [\pi_{i,t-1}; \widetilde{\pi}_{it}; \widetilde{y}_{it}]'$ and the vector of order $k_i^* \times 1$ of foreign variables $\mathbf{x}^*_{it} = [\pi^*_{i,t-1}; \widetilde{\pi}^*_{it}; \widetilde{y}^*_{it}]'$, and hence a GVAR version of the HNKPC is:

$$\mathbf{x}_{it} = \mathbf{a}_{i0} + \mathbf{\Phi}_i \mathbf{x}_{i,t-1} + \mathbf{\Lambda}_{i0}\mathbf{x}^*_{it} + \boldsymbol{\varepsilon}_{it}, \tag{6}$$

where $\mathbf{a}_{i0}$ is a $k_i \times 1$ vector containing constants to be estimated, Φ_i is a $k_i \times k_i$ matrix containing lagged coefficients, Λ_{i0} is a $k_i \times k_i^*$ matrix containing the foreign variables relevant for the country i, and $\boldsymbol{\varepsilon}_{it}$ is $k_i \times 1$ vector of errors. Notice that equation (6) could include more lags of the foreign variables vector, and it nests the VAR(1) if $\Lambda_{i0}{=}...{=}\Lambda_{ip^*}{=}0$. It is assumed that $\boldsymbol{\varepsilon}_{it} \sim iid(\mathbf{0}, \mathbf{\Sigma}_{ii})$; hence, errors are uncorrelated and with mean equal to 0. Note that $\mathbf{\Sigma}_{ii} = \mathbb{C}\left[\varepsilon_{ilt}, \varepsilon_{ist}\right]$ with $l \neq s$, and $\mathbf{\Sigma}_{ii}$ is nonsingular. This assumption could be easily relaxed for a spillover analysis with a long enough sample, since the elements of the diagonal must be estimated now. However, since $\mathbf{x}^*_{it}$ is included in the estimation, $\boldsymbol{\varepsilon}_{it}$ already contains some foreign information.

The foreign variables included in $\mathbf{x}^*_{it} = [\pi^*_{i,t-1}; \widetilde{\pi}^*_{it}; \widetilde{y}^*_{it}]'$ constitute a weighted average of the same variable defined for the remaining N countries:

$$\pi^*_{it} = \sum_{j=0}^{N} \omega^{\pi}_{ij}\pi_{jt}, \;\; \widetilde{\pi}^*_{it} = \sum_{j=0}^{N} \omega^{\widetilde{\pi}}_{ij}\widetilde{\pi}_{jt}, \;\; \widetilde{y}^*_{it} = \sum_{j=0}^{N} \omega^{\widetilde{y}}_{ij}\widetilde{y}_{jt}, \tag{7}$$

where $\{\{\omega^{\pi}_{ij}\},\{\omega^{\widetilde{\pi}}_{ij}\},\{\omega^{\widetilde{y}}_{ij}\}\}_{j=0}^{N}$ is the set of N weights for each of the k_i^* foreign variables relevant for the country i. The simplest weight scheme is the equally-weighted average with $\omega^{\pi}_{ij}{=}\omega^{\widetilde{\pi}}_{ij}{=}\omega^{\widetilde{y}}_{ij}{=}1/N$, $\forall i \neq j$. Obviously, as the sequences $\{\omega^{x}_{ij}\}$ are weights, $\Sigma_{j=0}^{N}\omega^{x}_{ij} = 1$.

A special attention is devoted to weights estimation in Gross (2013)'s article. A major claim by the author is that it is convenient to estimate them within the GVAR ensemble. This is because typically-used trade weights differ from those estimated, allowing for a chance to have a biased estimation of the GVAR parameters. The author also argues that weights leading to unbiased estimators may result in a better prediction performance. In this article, and according to the information extracted from a global inflation factor suggested in Ciccarelli and Mojon (2010), the weights coming from the first principal component are used when considering the set of six domestic inflation rates. This method also ensures to give an *ad hoc* weight to explain the majority of the whole set variance.

By now, equation (6) represents a VARX*(1,1) model, *i.e.* a VAR(1) model including exogenous variables X*. So, the advantage of the GVAR method is that it actually models all the variables contained in the weighted average. Hence, it includes the $N{+}1$ variables $\mathbf{x}_{it}$. This is made by stacking all the countries into one equation using the predetermined weights. As the weights are known, it is possible to estimate the equations separately and then continue with the stacking step.

Define the next $(k_i + k_i^*) \times 1$ vector $\mathbf{z}_{it}$:

$$\mathbf{z}_{it} = \begin{bmatrix} \mathbf{x}_{it} \\ \mathbf{x}^*_{it} \end{bmatrix}. \tag{8}$$

Equation (6) could be rewritten as:

$$\mathbf{A}_i \mathbf{z}_{it} = \mathbf{a}_{i0} + \mathbf{B}_i \mathbf{z}_{i,t-1} + \boldsymbol{\varepsilon}_{it}, \tag{9}$$

where $\mathbf{A}_i$ contains contemporaneous restrictions, $\mathbf{A}_i{=}[\mathbf{I}_k, -\mathbf{\Lambda}_{i0}]$, with $\mathsf{rank}(\mathbf{A}_i){=}k_i$ and $\mathbf{B}_i{=}[\mathbf{\Phi}_i, \mathbf{0}]$. If the foreign variables are included with a lag, then its coefficient matrix $\Lambda_{i,t-1}$, will appear in $\mathbf{B}_i$ as

$\mathbf{B}_i$=$[\mathbf{\Phi}_i, \mathbf{\Lambda}_{i,t-1}]$. A global vector $\mathbf{x}_t$ (suppressing the i-index) will be of the shape $\mathbf{x}_t = [\mathbf{x}_{0t}, \mathbf{x}_{1t}, ..., \mathbf{x}_{Nt}]'$, and the order in which the foreign variables enter $\mathbf{x}_{it}$ and the stacking order is irrelevant. To have a view on the matrices involved, let us have a look at the $\mathbf{A}_i$ shape for the case considered in this article:

$$\mathbf{A}_i = \begin{bmatrix} 1 & 0 & 0 & -\gamma_{ii}^{\widetilde{y}^*} & 0 & 0 \\ 0 & 1 & 0 & 0 & -\lambda_{ii}^{\pi^*} & 0 \\ 0 & 0 & 1 & 0 & 0 & -\lambda_{ii}^{\widetilde{\pi}^*} \end{bmatrix}. \tag{10}$$

Now, once that all the $\mathbf{x}_{it}$ vectors are already contained in the $\mathbf{z}_{it}$ vectors, it is easy to notice the following identity:

$$\mathbf{z}_{it} = \mathbf{W}_i \mathbf{x}_t, \tag{11}$$

where $\mathbf{W}_i$ (time-fixed) is a $(k_i + k_i^*) \times k$ matrix containing the known country-level weights. Pesaran, Schuermann, and Weiner (2004) label equation (11) as "the link", as it links the country-specific model $(\mathbf{z}_{it})$ using all the global variables $(\mathbf{x}_t)$. The shape of the $\mathbf{W}_i$ matrix when i=0 is shown below:

$$\mathbf{W}_{i=0} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_{01}^{\widetilde{y}} & 0 & 0 & \omega_{02}^{\widetilde{y}} & 0 & 0 & \omega_{03}^{\widetilde{y}} & 0 & 0 & \omega_{04}^{\widetilde{y}} & 0 & 0 & \omega_{05}^{\widetilde{y}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \omega_{01}^{\pi} & 0 & 0 & \omega_{02}^{\pi} & 0 & 0 & \omega_{03}^{\pi} & 0 & 0 & \omega_{04}^{\pi} & 0 & 0 & \omega_{05}^{\pi} & 0 \\ 0 & 0 & 0 & 0 & 0 & \omega_{01}^{\widetilde{\pi}} & 0 & 0 & \omega_{02}^{\widetilde{\pi}} & 0 & 0 & \omega_{03}^{\widetilde{\pi}} & 0 & 0 & \omega_{04}^{\widetilde{\pi}} & 0 & 0 & \omega_{05}^{\widetilde{\pi}} \end{bmatrix},$$

and the 3×3 submatrix of zeros (below the 3×3 identity submatrix) is moving one block (of 3 columns) to the right when the country is changed across i=1,...,5.

Using the link equation in the country-specific model delivers:

$$\mathbf{A}_i \underbrace{\mathbf{W}_i \mathbf{x}_t}_{\mathbf{z}_{it}} = \mathbf{a}_{i0} + \mathbf{B}_i \underbrace{\mathbf{W}_i \mathbf{x}_{i,t-1}}_{\mathbf{z}_{i,t-1}} + \varepsilon_{it}, \tag{13}$$

and $\mathbf{A}_i\mathbf{W}_i$ and $\mathbf{B}_i\mathbf{W}_i$ are both $k_i \times k$ matrices. Stacking these equations yields:

$$\mathbf{G}\mathbf{x}_t = \mathbf{a}_0 + \mathbf{H}\mathbf{x}_{t-1} + \varepsilon_t, \tag{14}$$

where:

$$\mathbf{a}_0 = \begin{bmatrix} \mathbf{a}_{00} \\ \mathbf{a}_{10} \\ \vdots \\ \mathbf{a}_{N0} \end{bmatrix}, \; \mathbf{G} = \begin{bmatrix} \mathbf{A}_0\mathbf{W}_0 \\ \mathbf{A}_1\mathbf{W}_1 \\ \vdots \\ \mathbf{A}_N\mathbf{W}_N \end{bmatrix}, \; \mathbf{H} = \begin{bmatrix} \mathbf{B}_0\mathbf{W}_0 \\ \mathbf{B}_1\mathbf{W}_1 \\ \vdots \\ \mathbf{B}_N\mathbf{W}_N \end{bmatrix}, \; \varepsilon_t = \begin{bmatrix} \varepsilon_{0t} \\ \varepsilon_{1t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}. \tag{15}$$

As $\mathbf{G}$ is a $k \times k$ matrix and of full rank generally, it is nonsingular allowing the GVAR representation:

$$\mathbf{x}_t = \mathbf{G}^{-1}\mathbf{a}_0 + \mathbf{G}^{-1}\mathbf{H}\mathbf{x}_{t-1} + \mathbf{G}^{-1}\varepsilon_t, \tag{16}$$

which can be solved recursively as a Structural VAR(1) model. Note that the structure of the model is commanded by the $\mathbf{G}$ matrix, which contains no row-crossed terms. This allows to estimate each country-level equation separately, to then stack all the $\mathbf{A}_i\mathbf{W}_i$ results (numerically) in $\mathbf{G}$. This method provides the advantage of achieving a large number of countries (or regions) and allowing different specifications for each country.

Many results are obtained from the estimation of equation (16). For the particular purpose of this article, I report the point estimate across the evaluation window of the lagged inflation coefficient, mimicking in a dynamic way the *persistence profile* suggested in Dees *et al.* (2007a, 2007b).

3.3 Forecast combinations

A traditional feature in forecasting literature is related to reaching accuracy improvements keeping fixed the information set. This task is typically explored through the so-called *forecast combinations*, launched after Bates and Granger (1969)'s article. Forecast combination relates simply to weight forecast of the same target variable at different horizons, opening a wide range of possibilities as the size of related literature attest. A particular case of combinations is that delivering the *combination puzzle* (Stock and Watson, 2004), *i.e.* simple weight-estimation procedures often outperform those obtained with fuzzy methods. The simplest method is to impose an equally-weighted scheme across the candidate forecasts, and is the alternative used in this article.[8]

In this article three combination schemes are used; C1: the RW combined with the CE-HNKPC, C2: the RW combined with the OE-HNKPC, and C3: both HNKPC. The former two combinations have the RW as a common element given that its accuracy deserves special attention when forecasting inflation, but also because it provides unbiased forecasts, as shown in Medel and Pincheira (2015). For a matter of exposition, consider iterating forward a general AR(1) model $y_t = \widetilde{c} + \phi y_{t-1} + \upsilon_t$, where υ_t is white noise and $\widetilde{c} = c(1-\phi)$, obtaining:

$$y_{t+h} = \widetilde{c}\left[\frac{1-\phi^h}{1-\phi}\right] + \phi^h y_t + \sum_{i=0}^{h-1}\phi^i \upsilon_{t+h-i}. \tag{17}$$

If y_t were a driftless RW ($\phi = 1$) then the optimal forecast would be y_t at any horizon. Accordingly, the expected value of the RW associated forecast error h-step-ahead forecast, $\mathbb{E}\left[\upsilon_t^{RW}(h)\right] = \mathbb{E}\left[y_{t+h} - y_t^{RW}(h)\right]$, would satisfy (see Medel and Pincheira, 2015, p. 127):

$$\begin{aligned} \mathsf{Bias}(h) &\equiv \mathbb{E}\left[\upsilon_t^{RW}(h)\right] = \mathbb{E}\left[c\left[\frac{1-\phi^h}{1-\phi}\right] - (1-\phi^h)y_t + \sum_{i=0}^{h-1}\phi^i_{t+h-i}\upsilon\right], \\ &= c\left[\frac{1-\phi^h}{1-\phi}\right] - (1-\phi^h)\mathbb{E}\left[y_t\right] = 0, \end{aligned} \tag{18}$$

as $\mathbb{E}\left[y_t\right] = c/(1-\phi)$; hence, becoming an unbiased forecast.

The C3 forecast is useful since it contains both economics-based models and will then be compared to a time series model; hence, evaluating the role of economic theory behind the HNKPC. The combined forecasts are then obtained according to:

$$\begin{aligned} \mathsf{C1} &: \pi^{\mathsf{C1}}_{t+h|t} = 0.5\pi^{\mathsf{RW}}_{t+h|t} + 0.5\pi^{\mathsf{CE\text{-}HNKPC}}_{t+h|t}, \\ \mathsf{C2} &: \pi^{\mathsf{C2}}_{t+h|t} = 0.5\pi^{\mathsf{RW}}_{t+h|t} + 0.5\pi^{\mathsf{OE\text{-}HNKPC}}_{t+h|t}, \\ \mathsf{C3} &: \pi^{\mathsf{C3}}_{t+h|t} = 0.5\pi^{\mathsf{CE\text{-}HNKPC}}_{t+h|t} + 0.5\pi^{\mathsf{OE\text{-}HNKPC}}_{t+h|t}. \end{aligned} \tag{19}$$

[8] Note that robustness checks using a weighting scheme of (0.75;0.25) and (0.25;0.75) is also analysed (see equation 19). The results–details available upon request–indicate worst performance using the pair (0.25;0.75) for C1 at h=24, and C2 and C3 at h=12, when considering outliers. Without outliers, there are improvements noticed for C2 and C3 at h=24, exhibiting a RMSFE Ratio of 0.994 and 0.996, compared to the current 1.048 and 1.020, respectively–see Table 2. However, the differences between baseline and alternative weighting schemes are not statistically significant, supporting the *combination puzzle* claim.

3.4 Forecast evaluation framework

The statistical measure used to evaluate the accuracy of point forecasts is the RMSFE:

$$\mathsf{RMSFE}_h = \left[\frac{1}{T(h)} \sum_{t=1}^{T(h)} (\pi_{t+h|t+h} - \pi_{t+h|t})^2 \right]^{\frac{1}{2}}, \tag{20}$$

where $\pi_{t+h|t}$ is the h-step-ahead forecast of $\pi_{t+h|t+h}$ made at period t. Note that this statistic is computed given a forecasting horizon h, and hence, the difference $T-t$ is a variable depending on h ($T = T(h)$). To make a more plausible comparison with the RW, the analysed statistic corresponds to the RMSFE Ratio defined as:

$$\mathsf{RMSFE\ Ratio}_h = \frac{\mathsf{RMSFE}_h^{\mathcal{M}}}{\mathsf{RMSFE}_h^{\mathsf{RW}}}, \tag{21}$$

where $\mathcal{M}$={CE-HNKPC,OE-HNKPC,AR,ES,C1,C2,C3}. Hence, as the RW acts as a pivot, values greater than unity imply a worse performance of the competing model. Figures below unity represent a "predictive gain" of (1-RMSFE Ratio)% upon the RW.

Note that this evaluation is specifically made by "country×variable" forecast elements (the identifier is unique). Nevertheless, from the GVAR it is possible to evaluate the predictive accuracy of all the variables comprising a single country, a region, or a set of variables (where the "country×variable" elements are no longer unique).

To investigate to what extent the predictive gains are statistically significant, I make use of the unconditional t-type test of GW providing the advantage of comparing *forecasting methods* instead of *forecasting models*. As the null hypothesis (NH) is defined as *the competing model that has a superior predictive ability compared to the* RW, there a one-side t-type GW statistic is used accordingly.

Formally, the NH: $\mathbb{E}_t(d_h) \leq 0$ is tested against the alternative AH: $\mathbb{E}_t(d_h) > 0$, where:

$$\mathsf{d}_h = (\pi_{t+h|t+h} - \pi_{t+h|t}^{\mathsf{RW}})^2 - (\pi_{t+h|t+h} - \pi_{t+h|t}^{\mathcal{M}})^2, \tag{22}$$

using the Newey and West (1987) HAC estimator of the standard deviation of d_h. The NH is rejected if the subsequent t-statistic is greater than $t_{\alpha\%}$; corresponding to the tabulated value of a normal distribution with probability $\alpha\%$.

4 Results

This section analyses three kinds of results: in-, out-of-sample estimates, and robustness exercises. The in-sample results are related to estimation diagnostics and stability, whereas the out-of-sample results exclusively to dynamic forecasts precision (RMSFE Ratio). Finally, robustness exercises are related to an open-economy version of the CE-HNKPC using RER information and a moving average transformation of the output gap.

4.1 In-sample diagnostics

This subsection primarily analyses the econometric diagnostic behind the estimation of the two economics-based models. Table 1 presents the coefficient estimation results of the CE-HNKPC using the estimation sample. Although the main focus is the Chilean economy, the results for the other economies are shown for reference. In particular, all these results are useful since they do not reject the hypothesis that estimates actually obeys to a Phillips Curve.

The results deliver similar estimations to that exhibited in Medel (2015b) when comparable. Moreover, the results shown in this article are closer to that dictated by the theory. Note that the dependent variable, in this case, is the difference between actual inflation and the inflation target of each country: $\widehat{\pi}_t = \pi_t - \pi^{Target}$. The level of confidence in which all models are statistically significant is 15%. In particular, with valid IV as suggested by the J-statistic p-value (0.181), the coefficient of the output gap is positive and statistically significant. Also, all these estimations are done without any restriction, in particular, without imposing $\lambda_b + \lambda_f$=1. Nevertheless, the sum of both mentioned parameters achieve 1.351 and a ratio of λ_f/λ_b=0.75. This imply that, when re-scaled to add unity, the parameters are λ_b=57% and λ_f=43%. The adjusted goodness-of-fit coefficient suggests that the model potentially has a good predictive power, and it is well specified according to the DW statistic. However, as Hansen (2009) argues, it is not clear the relationship between in-sample fit and forecast accuracy, but forecasts tend to be worst with overfitted models.[9]

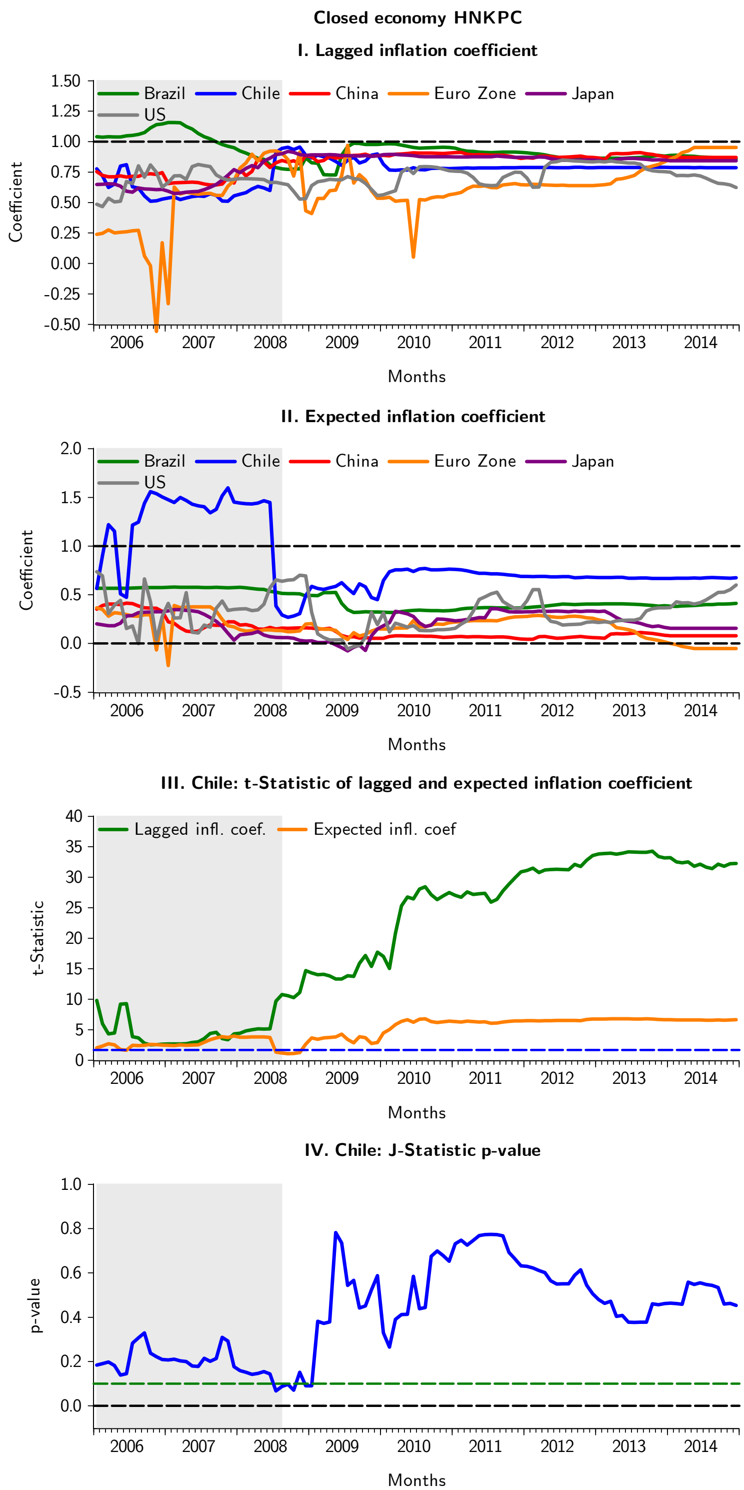
Table 1: GMM estimates of the HNKPC (*)

	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*
	Dependent variable: $\widehat{\pi}_t = \pi_t - \pi^{Target}$					
	Estimation sample					
π^{Target}	4.5%	3.0%	-	2.0%	2.0%	2.0%
$\widehat{\pi}_{t-1}$	1.039	0.773	0.766	0.248	0.690	0.509
	[0.000]	[0.000]	[0.000]	[0.045]	[0.000]	[0.020]
$\widetilde{\pi}_t$	0.565	0.578	0.346	0.349	0.194	0.737
	[0.000]	[0.047]	[0.013]	[0.020]	[0.006]	[0.000]
$\widetilde{y}_t$	0.776	0.072	0.211	0.039	0.019	0.190
	[0.124]	[0.036]	[0.091]	[0.142]	[0.069]	[0.065]
$\overline{\pi}$	-3.979	-1.672	-0.265	-0.347	-0.736	-1.168
	[0.000]	[0.057]	[0.265]	[0.085]	[0.001]	[0.001]
$\overline{R}^2$	0.843	0.845	0.721	0.267	0.702	0.653
S.E. Reg	1.307	0.479	0.871	0.233	0.240	0.483
DW Stat.	0.803	1.818	1.818	1.175	1.557	1.525
J-Stat.	0.591	4.873	2.308	5.494	3.422	5.326
p-value	0.743	0.181	0.679	0.240	0.180	0.149
	Instrumental variables list (lags)					
Constant	✓	✓	✓	✓	✓	✓
π_t	(2), (6)	(8)	(2), (5)	(2), (6)	(2)	(2)
$\widetilde{\pi}_t$	(1), (4)	(1), (5)	(2)	(8)	(2)	(1), (9)
$\widetilde{\pi}_t^{Current}$	-	-	(7)	(4)	-	-
$\widetilde{\pi}_t^{Next}$	-	-	(9)	(7)	-	-
$\widetilde{y}_t$	(6)	(1), (4), (7)	(2), (11)	(4), (8)	(4), (11), (12)	(1), (3), (9)

(*) Equation: $\widehat{\pi}_t = \overline{\pi} + \lambda_b\widehat{\pi}_{t-1} + \lambda_f\widetilde{\pi}_t + \gamma\widetilde{y}_t + \varepsilon_t$, with $\varepsilon_t \sim iid\mathcal{N}(0, \sigma^2_\varepsilon)$. For China, the $\widetilde{\pi}_t$ variable corresponds to $\widetilde{\pi}^{Current}_{t-p}$. Coefficient p-value in $[\cdot]$. "DW Stat." stand for the Durbin-Watson statistic. Source: Author's elaboration.

[9] See Medel (2015c) for some calibrated estimations of the effect of overfitting in the quality of the predictions, and Calhoun (2014) for a theoretical background.

Figure 1: Closed economy HNKPC: recursive estimates
of lagged and expected inflation coefficients and inference (*)



(*) Shaded area = shortened evaluation sample. Horizontal line in III = 1.65. Horizontal line in IV = 10%. Source: Author's elaboration.

Note that the convergence to this specification, and particularly the IV lags, is made using a *General-to-Specific* procedure matching not only joint and individual significance, but IV appropriateness also. It is worth mentioning that to find a specification that fulfils all desirable statistical and economic checks is a daunting task. This is a recognised problem with this kind of estimations, which redound in a particularly unstable environment. These results for are not robust to changes in the IV sets. Therefore, further results should be taken with caution. Note that even the use of a richer structural model does not necessarily redound in a more stable estimation or a robust calibration. This is due to the difficulty to match the moments of a set of variables containing, for instance, empirical puzzles.

Precisely with the aim of analysing instability, in Figure 1 I present a recursive estimation across the evaluation sample of several key parameters of the model. Panel I depicts the coefficient of the lagged inflation ($\widehat{\pi}_t$) for all the countries considered. The results show an astonishing stable result for Brazil, whereas for Chile and the Euro Zone there are major disturbances during the 2008-9 financial crisis. For the remaining countries (China, Japan, and the US), the estimations start to be stable in 2010. All these parameters are statistically significant with the estimation sample. The behaviour of this coefficient for Chile is not surprising. In line with Figure B1, Chile exhibits a major inflation peak during the mentioned episode. For Brazil it is also easy to notice a high inflation period but located in the estimation sample (2003), and showing no major reaction to the 2008-9 disturbances.

Another component of inflation persistence is the coefficient of expected inflation, which is depicted in a recursive manner in Panel II. In this case, same dynamics are roughly observed for Brazil, China, the Euro Zone, Japan, and the US, similar to the previous case. Remarkably, for the Chilean case the expected inflation coefficient achieves 1.5 in the beginning of the (evaluation) sample. When the financial crisis hit, the parameter fell to then stay steady since 2010 onwards. Moreover, while the lagged inflation coefficient grew, the expected coefficient fell down at the same time. This dynamic is of particular interest since the parameters are not restricted to adding to a constant, although the estimates behave as if they already are. This fact also suggests that the model is capturing well the mechanics of the HNKPC, and that the inflation expectations variable is a valid measure.

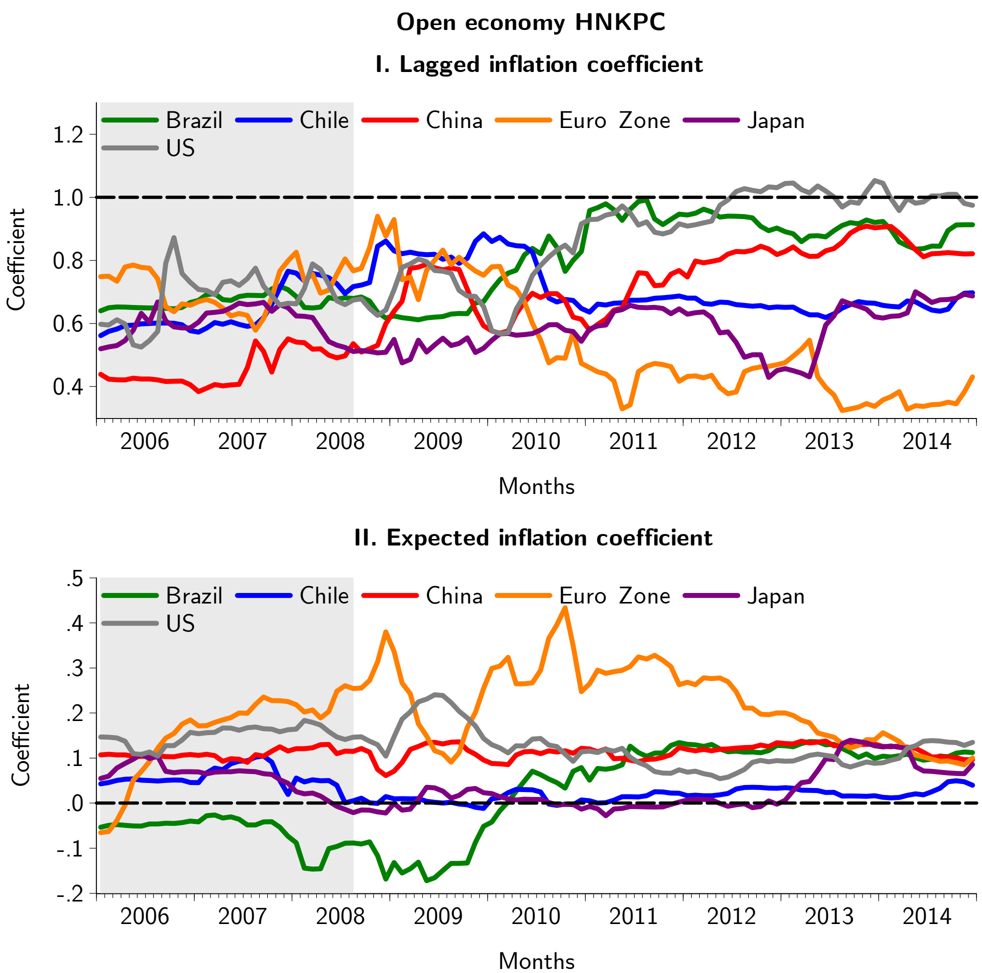
Panels III and IV show statistical inference just for the Chilean case. The former depicts the t-statistic of both the lagged and expected inflation coefficients while the latter shows the J-statistic p-value of IV validity. Note that the IV specification that feeds the second stage estimation is valid most of the time. Regarding the significance of the coefficients, the lagged inflation coefficient is always significant. The expected inflation coefficient loses its significance during the 2008-9 period, although then it is recovered and always positioned above the 95% confidence level threshold. In sum, it is concluded that the CE-HNKPC for the Chilean economy has a robust estimation.

A slightly different picture is obtained with the OE-HNKPC. Figure 2 presents the same two coefficient estimates shown for the CE-HNKPC, *i.e.* the lagged and expected inflation coefficients. From Panel I, it is possible to notice that all the estimations lie in the (0.3,1.0) range across the sample–different to the previous case–but with a remarkable more volatile dynamic. Chile results in the most stable estimation, which is not a surprising result since it is the country that actually includes information from its bigger trading partners. Brazil is the second country in representativeness of its major trading partners. There are noticeably two periods in its coefficient dynamics, *i.e.* before and after the financial crisis.

Panel II depicts the recursive estimation of the expected inflation coefficient. In this case more stable coefficients are observed compared to the CE-HNKPC. However, major differences are found for the Euro Zone and to a lesser extent Brazil, showing again a two-regime-alike estimates. For Chile the results are stable but close to zero during the 2008-9 period.

More econometric diagnostics of the GVAR are presented in Appendix D. Particularly, the residuals of the 18 equations are depicted over time. All the residuals are well behaved exhibiting the required white noise behaviour. For all the countries, except China, the equation of expected inflation contains at least one outlier (which does not deserve any correction) in different periods of time. Finally, autoregression diagnostics are presented in Appendix E.

Figure 2: Open economy HNKPC: recursive estimates
of lagged and expected inflation coefficients (*)



(*) Shaded area = shortened evaluation sample. Source: Author's elaboration.

4.2 Out-of-sample results

This subsection presents the out-of-sample results for both the evaluation and shortened sample. These results comprise the RMSFE Ratio of equation 21, and are presented in Table 2. Note that all forecast are made for the π_t variable.

In the shortened sample, the AR model is the best alternative for the most immediate horizon, followed by the combination between the RW and the OE-HNKPC (C2), and both economics-based models (C3), noting that none of these superiority results are statistically significant. For h=6 none of the proposed models are superior to the RW. At h=12, the AR model again plus the CE-HNKPC and its combination with the RW is better than the RW alone. Despite that the best adjustment is found to the AR model, the C1 forecast results in a statistically significant superiority. In the long-run, the best alternative is the C2 forecast but not resulting statistically superior. It is hence obtained that, using the shortened sample, the best options are the AR plus either of the two combined forecasts, giving not a clear role for economics-based models.

Table 2: Chile: RMSFE Ratio estimates (*)

	AR	ES	CE-HNKPC	OE-HNKPC	C1	C2	C3	RW
				Shortened sample				(RMSFE)
h=1	0.934	1.568⋆	1.074	1.082	1.016	0.946	0.972	0.493
h=6	1.114	1.100⋆	1.070⋆	1.284	1.022	1.116⋆	1.141⋆	2.194
h=12	0.970	1.001	0.983	1.044	0.987⋆	1.017	1.006	3.926
h=24	1.187⋆	1.001	1.255	1.008⋆	1.127⋆	0.996	1.125⋆	4.579
				Evaluation sample				(RMSFE)
h=1	0.888⋆	1.487⋆	0.913	0.949	0.923⋆	0.886⋆	0.864⋆	0.625
h=6	0.934	1.088⋆	0.897	1.108 [0.986]	0.918	0.772⋆ [0.763⋆]	0.785 [0.769]	2.474
h=12	0.915⋆	1.015	1.022	1.552 [1.218]	0.986	0.886 [0.825]	0.960 [0.890]	4.037
h=24	0.831⋆	0.964⋆	1.053	2.017 [1.289]	0.983	1.344 [1.048]	1.291 [1.020]	4.490

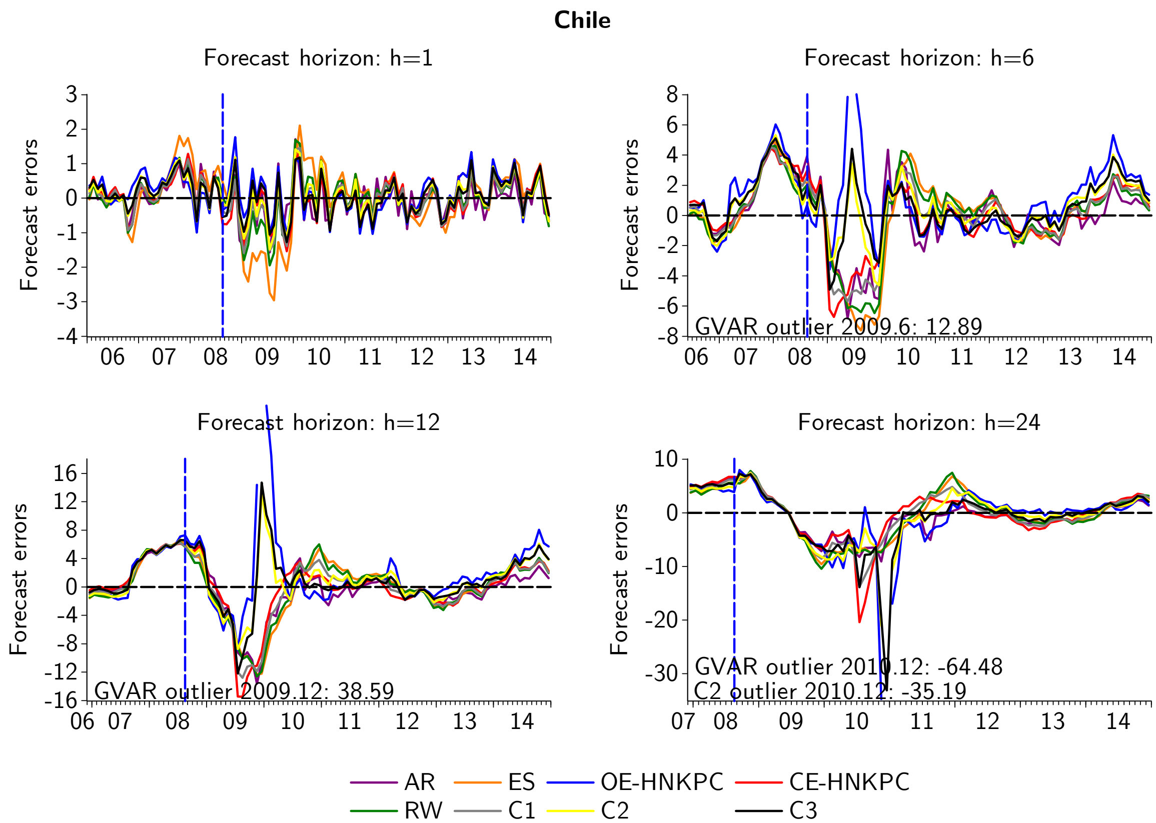
(*) Shaded cells = figures below unity (without outliers). GW test results: (⋆⋆⋆) p<1%, (⋆⋆) p<5%, (⋆) p<10%. RMSFE Ratios in [·] are computed without outliers. Source: Author's elaboration.

With the complete evaluation sample, more alternatives and results emerge. It is worth mentioning that the OE-HNKPC exhibit several outliers in forecasting error series when predicting at h={6,12,24}. The RMSFE Ratio results are presented for both series either containing or not mentioned observations. It is a valid option to drop these observations since they are already outliers, not following a systematic pattern. Also, the sample size with which the RMSFE are calculated is long enough to give a minor weight to a particular observation despite its size. The result of Table 2 indicates that in just one case (OE-HNKPC; h=6) the outlier correction changes the meaning of the results, *i.e.* lowering the ratio from above to below unity.

The analysis is then conducted without outliers. Now, the three combined forecasts provide an overwhelmingly superior predictive ability up to the 12-month ahead horizon, and C1 also for h=24. More importantly, the combined HNKPC forecast (C3) itself provides superior results than the benchmark at short horizons (h={1,6}). Note that the AR outperforms the RW at any horizon, whereas the ES is statistically superior at h=24. Any combination scheme is not worse than the AR model at h=1. At h=6, only C2 outperforms the AR model. It is hence obtained that when considering the whole evaluation sample, characterised with an increase in targeted variable volatility, the economics-based models are superior in horizons within a year, while at longer horizons the best alternatives are statistical models.

To have an in-depth analysis of the dynamics of the forecasting errors, in Figure 3 the forecasting errors across the evaluation sample for all the models and horizons are depicted. This figure also points out the random character of dropped outliers. Also, this figure suggests that the worse tracking during the financial crisis is made with the ES model, while the best results are obtained with the CE-HNKPC and the C2 forecast. At h=6, there are two forecasts showing more precise results during the crisis, the C2 and C3 forecasts. The C2 contains information from the CE-HNKPC prediction (being valuable also for this horizon), whereas the C3 results as a valid option since the OE-HNKPC errors are offset by a downward error (overprediction) of the RW. At h=12, similar results to those with h=6 are obtained. For these both horizons, most of the forecasting error variance obviously comes from the unanticipated effect of the financial crisis. At h=24, Figure 3 shows that the statistical models are all near to each other whereas the economics-based models exhibit several peaks during the 2010-11 period. Note, however, that in normal times all the models behave similarly.

Figure 3: Chile: Multihorizon forecasting errors across evaluation sample (*)



(*) Vertical line = shortened evaluation sample end point. Source: Author's elaboration.

4.3 Robustness exercises

In this subsection two alternative specifications are analysed for the CE-HNKPC as robustness check. These are in line with the traditional view found in the literature regarding an open-economy version of the CE-HNKPC. As abovementioned, there is neither a unique nor consensued way in which a close-economy HNKPC could be transformed into an open-economy version. However, as the aim of this article is inflation forecast accuracy through a HNKPC ensemble, parsimonious models are always preferred. Hence, the first check is to analyse if the RER dynamics–the annual percentage change of the RER index, q_t–plays a significant role once included in the baseline specification. The inclusion of RER made in this manner obeys to the simplest specification.

It is worth mentioning that an open-economy specification could involve fuzzy specifications, particularly in the construction of the output gap (see, for instance, Posch and Rumler, 2014). Nevertheless, complicated specifications are often associated with a larger number of variables and parameters, to which auxiliary forecast are necessary. In the case of Posch and Rumler (2014), for instance, an open-economy output gap specification lies also in steady-state shares of labour, domestic intermediate inputs, and imported intermediate inputs in total domestic production. Then, an AR model is used to predict the resulting output gap required for the inflation forecast, avoiding, to some extent, the economic content that proposed variables may provide. In sum, the inclusion of a richer structure may work for (in-sample) testing the economic theory behind the model, whereas the task of (out-of-sample) forecasting lies conveniently in statistical modelling.

Despite that the model is specified for a closed economy, the actual inflation data is permeable to foreign components, presumably higher in countries with a larger trade-based sector, such as Japan. This fact reflects that data already contains foreign-countries information. The specification search is made in the same manner as before, that is, iterating through different lags acting as IV, aiming to match individual significance as well as IV validity. The results are presented in Table 3. Note that the results, in line with previous finding of Medel (2015b) for the Chilean case, indicate that RER either is non statistically significant or spoils the baseline specification. Hence, neither of these specifications are used to forecast. Remarkably, Lubik and Schorfheide (2007)–analysing the reaction of four central banks of industrialised countries to foreign variables such as nominal exchange rate–find that terms-of-trade do not contribute significantly to domestic business cycles.

Table 3: GMM estimates of the HNKPC including RER (*)

	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*
	Dependent variable: $\widehat{\pi}_t = \pi_t - \pi^{Target}$					
	Estimation sample					
π^{Target}	4.5%	3.0%	-	2.0%	2.0%	2.0%
$\widehat{\pi}_{t-1}$	0.874	0.685	0.674	0.853	0.546	0.497
	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.023]
$\widetilde{\pi}_t$	0.317	0.925	0.536	0.387	0.282	0.758
	[0.022]	[0.019]	[0.000]	[0.037]	[0.097]	[0.000]
$\widetilde{y}_t$	0.166	0.065	0.030	-0.113	0.037	0.191
	[0.510]	[0.039]	[0.627]	[0.024]	[0.012]	[0.071]
q_t	-0.024	-0.004	0.088	0.007	-0.021	0.00085
	[0.092]	[0.628]	[0.004]	[0.309]	[0.005]	[0.928]
$\overline{\pi}$	-1.440	-2.750	-0.382	-0.651	-1.106	-1.201
	[0.004]	[0.022]	[0.008]	[0.027]	[0.004]	[0.002]
$\overline{R}^2$	0.966	0.849	0.873	-0.113	0.654	0.643
S.E. Reg	0.602	0.473	0.587	0.288	0.259	0.490
DW Stat.	1.144	1.749	1.480	1.232	1.259	1.512
J-Stat.	2.467	4.773	5.854	5.299	2.056	5.283
p-value	0.481	0.311	0.210	0.257	0.151	0.152
	Instrumental variables list (lags)					
Constant	✓	✓	✓	✓	✓	✓
π_t	(2), (6)	(8)	(2), (5)	(2), (6)	(2)	(2)
$\widetilde{\pi}_t$	(1), (4)	(1), (5)	(5)	(8)	(2)	(1), (9)
$\widetilde{\pi}_t^{Current}$	-	-	(9)	(4)	-	-
$\widetilde{\pi}_t^{Next}$	-	-	(7)	(7)	-	-
$\widetilde{y}_t$	(6)	(1), (4), (7)	(2), (11)	(4), (8)	(4), (12)	(1), (3), (9)
q_t	(1), (3)	(1), (5)	(1)	(12)	(1)	(1)

(*) Equation: $\widehat{\pi}_t = \overline{\pi} + \lambda_b \widehat{\pi}_{t-1} + \lambda_f \widetilde{\pi}_t + \gamma \widetilde{y}_t + \kappa q_t + \varepsilon_t$, with $\varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$.
See notes to Table 1. Source: Author's elaboration.

A second robustness check consists in the use of another statistical specification of the output gap. Particularly, a 12-order moving average version of the output gap substituting the baseline specification of models in Table 1 is used. The aim of this exercise is to analyse if inner movements of output gap–as the moving average captures–are still related to current values of inflation.

Table 4: GMM estimates of the HNKPC using MA output gap (*)

	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*
	Dependent variable: $\widehat{\pi}_t = \pi_t - \pi^{Target}$					
	Estimation sample					
π^{Target}	4.5%	3.0%	-	2.0%	2.0%	2.0%
$\widehat{\pi}_{t-1}$	0.795	0.726	0.842	0.268	0.479	0.653
	[0.000]	[0.000]	[0.000]	[0.050]	[0.032]	[0.000]
$\widetilde{\pi}_t$	0.520	0.396	0.254	0.255	0.583	0.304
	[0.000]	[0.118]	[0.000]	[0.113]	[0.026]	[0.109]
$\widetilde{y}_t^{MA}$	0.286	0.095	0.257	0.047	0.029	0.089
	[0.000]	[0.007]	[0.135]	[0.112]	[0.075]	[0.082]
$\overline{\pi}$	-2.531	-1.214	-0.073	-0.188	-0.999	-0.412
	[0.000]	[0.112]	[0.487]	[0.366]	[0.032]	[0.259]
$\overline{R}^2$	0.972	0.878	0.860	0.343	0.551	0.793
S.E. Reg	0.564	0.423	0.628	0.220	0.307	0.372
DW Stat.	1.196	1.333	1.829	1.282	0.857	1.296
J-Stat.	0.622	2.997	4.941	6.350	4.499	3.357
p-value	0.732	0.223	0.176	0.174	0.105	0.186
	Instrumental variables list (lags)					
$Constant$	✓	✓	✓	✓	✓	✓
π_t	(2), (6)	(10)	(2), (5)	(2), (7)	(2)	(3)
$\widetilde{\pi}_t$	(1), (4)	(1), (5)	(5)	(5), (7)	(2)	(1)
$\widetilde{\pi}_t^{Current}$	-	-	(9)	(12)	-	-
$\widetilde{\pi}_t^{Next}$	-	-	(7)	-	-	-
y_t	-	-	-	(2)	(1), (9)	(10)
$\widetilde{y}_t^{MA}$	(6)	(1), (2)	(10)	(1)	(5)	(1), (2)

(*) Equation: $\widehat{\pi}_t = \overline{\pi} + \lambda_b \widehat{\pi}_{t-1} + \lambda_f \widetilde{\pi}_t + \gamma \widetilde{y}_t^{MA} + \varepsilon_t$, with $\varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$.
See notes to Table 1. For China and Japan the $\widetilde{\pi}_t$ variable corresponds to $\widetilde{\pi}_{t-p}^{Current}$. Source: Author's elaboration.

The results are presented in Table 4. It is observed that the CE-HNKPC hypothesis is still not rejected for all the countries considered. Interestingly, estimates of the remaining parameters are closer to that found in the baseline specification. As models of Table 3 fulfil the desirable economic and statistical requirements, they are used for forecasting. The results of this task are analysed in comparative terms following the RMSFE Robustness Ratio:

$$\text{RMSFE Robustness Ratio}_h = \frac{\text{RMSFE}_h^{Baseline}}{\text{RMSFE}_h^{MovingAverage}}, \tag{23}$$

where a ratio below unity indicates that the baseline is more accurate than the moving average specification. These results are displayed in Table 4 for all the countries. Note that precisely for Chile there are three cases observed in which the baseline model is outperformed (at h={1,6,12}). However, these predictive gains are not superior to 3% and not statistically significant. Other favourable cases for the moving-average output gap are China (0.9%, h={1}) and the Euro Zone (1.6%, h={12}), but also of negligible size. It is concluded, hence, that for the Chilean case the alternative output gap measure already plays a role in forecast accuracy, however not overwhelmingly superior to that of the baseline specification. In economic terms, it is suggested that the persistent dynamics of economic slack is also a determinant of current inflation.

Table 5: RMSFE Ratio between baseline and moving average gap specification (*)

	Evaluation sample					
	CE-HNKPC (Baseline/Moving average)					
	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*
h=1	0.923	1.004	1.009	0.943	0.991	0.991
h=6	0.985	1.025	0.859	0.920	0.939	0.937
h=12	0.993	1.032	0.688	1.016	0.856	0.964
h=24	0.418	0.977	0.437	0.845	0.775	0.860

	OE-HNKPC (Baseline/Moving average)					
	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*
h=1	0.820	0.970	0.991	0.850	0.959	0.839
h=6	0.676	0.765	0.955	0.669	0.968	0.696
h=12	0.519	0.767	0.869	0.653	1.245	0.649
h=24	0.851	0.452	0.458	1.113	1.222	1.256

(*) Shaded cells = figures below unity.

Source: Author's elaboration.

5 Summary and concluding remarks

This article has analysed the multihorizon predictive power of the HNKPC for the Chilean inflation, making use of closed- and open-economy versions (CE-HNKPC and OE-HNKPC); the latter based in a GVAR ensemble including the Chilean main trade partners, namely Brazil, China, the Euro Zone, Japan, and the US, completing up to 70% of its total trade.

These economics-based forecasts are compared with traditional time-series benchmarks used in the literature, plus three combined forecasts, leaving also the option to evaluate the isolated economic content of the HNKPC in an out-of-sample context. The analysed monthly sample covers from 2000.1 to 2014.12, divided into the estimation sample (2000.1-2005.12) and the evaluation sample (2006.1-2014.12). A special focus is given to the period 2006.1-2008.8 (just before the financial crisis); hence, evaluating it in *normal* times too. The analysed forecast horizons are h={1,6,12,24} months ahead.

The driving process of the HNKPC in its two versions is the Hodrick-Prescott-based output gap with a treatment for the end-of-sample problem; similar to that used in Medel (2015a, 2015b). One of the key elements of this article is the use of direct measures of inflation expectations embedded in the two versions of the HNKPC for forecasting purposes–and different from the case where inflation expectations are computed within the model. The expectations are taken from the monthly *Consensus Forecasts* report, and transforming to a unique variable given its fixed-horizon nature. The HNKPC is robust to a moving average output gap specification, suggesting that persistent economic slack is a determinant of current inflation values.

The results indicate that there is evidence to do not reject the hypothesis of the HNKPC for the Chilean economy, *i.e.* that the lagged and expected inflation coefficients are statistically significant, as is also that of the output gap. This finding is obtained with the CE-HNKPC. The OE-HNKPC specification introduced in this article also complies with the required statistical and economic-based tests. In predictive terms, the out-of-sample results show that with the shortened sample the evidence is mixed between atheoretical statistical models and the HNKPC itself or in a combined prediction. However, when the evaluation sample is extended to a more volatile period, the results suggest that both versions of the HNKPC (and combined with the RW) deliver the most accurate forecasts at horizons comprised within a year.

In the long run, the combination between the CE-HNKPC and the RW delivers more accurate results than the benchmark, however not enough to outperform the statistical models. Note also that the results for the OE-HNKPC have to deal with outliers exhibited during the financial crisis, although not threatening the main conclusions. It is hence concluded that at short horizons and when inflation show higher volatility, the HNKPC results in the best forecasting option compared to traditional statistical models; a finding that is reverted at longer horizons.

References

1. Andrews, D.W.K., 1993, "Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models," *Econometrica* **61**(1): 139-165.

2. Bates, J.M. and C.W.J. Granger, 1969, "The Combination of Forecasts," *Operational Research Quarterly* **20**(4): 451-468.

3. Batini, N., B. Jackson, and S. Nickell, 2005, "An Open-Economy New Keynesian Phillips Curve for the UK," *Journal of Monetary Economics* **52**(6): 1061-1071.

4. Bobbitt, L. and M.C. Otto, 1990, "*Effects of Forecasts on the Revisions of Seasonally Adjusted Values Using the X-11 Seasonal Adjustment Procedure*," Proceedings of the Business and Economic Statistics Section, American Statistical Association, 449-453.

5. Box, G.E.P. and G.M. Jenkins, 1970, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, US.

6. Calhoun, G., 2014, "*Out-of-Sample Comparisons of Overfit Model*," Working Paper 11002, Department of Economics, Iowa State Univeristy, US.

7. Calvo, G.A., 1983, "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics* **12**(3): 383-398.

8. Canova, F., 2007, "G-7 Inflation Forecasts: Random Walk, Phillips Curve or What Else?" *Macroeconomic Dynamics* **11**: 1-30.

9. 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.

10. 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.

11. Ciccarelli, M. and B. Mojon, 2010, "Global Inflation," *Review of Economics and Statistics* **92**(3): 524-535.

12. Cochrane, J., 2001, *Asset Pricing*, Princeton University Press, US.

13. Dees, S., F. di Mauro, M.H. Pesaran, and L.V. Smith, 2007a, " Exploring the International Linkages of the Euro Area: A Global VAR Analysis," *Journal of Applied Econometrics* **22**(1): 1-38.

14. Dees, S., S. Holly, M.H. Pesaran, and L.V. Smith, 2007b, "Long Run Macroeconomic Relations in the Global Economy," *Economics: The Open-Access, Open-Assessment E-Journal* **3**(2007): 1-58.

15. Dees, S., M.H. Pesaran, L.V. Smith, and R.P. Smith, 2009, "Identification of New Keynesian Phillips Curves from a Global Perspective," *Journal of Money, Credit and Banking* **41**(7): 1481-1502.

16. Dickey, D.A. and S.G. Pantula, 1987, "Determining the Order of Differencing in Autoregressive Processes," *Journal of Business and Economic Statistics* **5**(4): 455-459.

17. Erceg, C.J. and A.T. Levin, 2003, "Imperfect Credibility and Inflation Persistence," *Journal of Monetary Economics* **50**(4): 915-944.

18. Faust, J. and J. Wright, 2014, *Forecasting Inflation*, in G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting, Volume 2*, Elsevier, North-Holland.

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

20. Galí, J., M. Gertler, and J.D. López-Salido, 2001, "European Inflation Dynamics," *European Economic Review* **45**(7): 1237-1270.

21. 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): 1107-1118.

22. Galí, J. and T. Monacelli, 2005, "Monetary Policy and Exchange Rate Volatility in a Small Open Economy," *Review of Economic Studies* **72**: 707-734.

23. Giacomini, R. and H. White, 2006, "Tests of Conditional Predictive Ability," *Econometrica* **74**(6): 1545-1578.

24. Ghysels E., D. Osborn, and P.M. Rodrígues, 2006, *Forecasting Seasonal Time Series*, in G. Elliott, C.W.J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, *Volume 1*, Elsevier, North Holland.

25. Granger, C.W.J. and Y. Jeon, 2011, "The Evolution of the Phillips Curve: A Modern Time Series Viewpoint," *Economica* **78**: 51-66.

26. Gross, M., 2013, "*Estimating GVAR Weight Matrices*," Working Paper 1523, European Central Bank.

27. Hansen, L.P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* **50**(4): 1029-1054.

28. Hansen, P.R., 2009, "*In-Sample Fit and Out-of-Sample Fit: Their Joint Distribution and its Implications for Model Selection,*" *manuscript*, version of 23 April, 2009, Department of Economics, Stanford University, US.

29. Henzel, S. and T. Wollmershaeuser, 2008, "The New Keynesian Phillips Curve and the Role of Expectations: Evidence from the CESifo World Economic Survey," *Economic Modelling* **25**(5): 811-832.

30. Hyndman, R.J., A.B. Koehler, J.K. Ord, and R.D. Snyder, 2008, *Forecasting with Exponential Smoothing. The State Space Approach*, Springer Series on Statistics, Berlin, Germany.

31. Jean-Baptiste, F., 2012, "Forecasting with the New Keynesian Phillips Curve: Evidence from Survey Data," *Economics Letters* **117**(3): 811-813.

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

33. Kullback, S. and R.A. Leibler, 1951, "On Information and Sufficiency," *Annals of Mathematical Statistics* **22**: 79-86.

34. Levin, A., A. Onatski, A. Williams, and J. Williams, 2005, *Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models*, in M. Gertler and K. Rogoff (Eds.), *NBER Macroeconomics Annual*, MIT Press, US.

35. Lindé, J., 2005, "Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach," *Journal of Monetary Economics* **52**(6): 1135-1149.

36. Lovell, M., 2008, "A Simple Proof of the FWL Theorem," *Journal of Economic Education* **39**(1): 88-91.

37. Lubik, T.A. and F. Schorfheide, 2007, "Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation," *Journal of Monetary Economics* **54**: 1069-1087.

38. Medel, C.A., 2015a, "*Forecasting Inflation with the Hybrid New Keynesian Phillips Curve: A Compact-Scale Global VAR Approach*," MPRA Paper 67081, University Library of Munich, Germany.

39. Medel, C.A., 2015b, "Inflation Dynamics and the Hybrid New Keynesian Phillips Curve: The Case of Chile," *Monetaria* **III**(1): 25-69.

40. Medel, C.A., 2015c, "Classical Probability of Overfitting with Information Criteria: Estimates with Chilean Macroeconomic Series," [in Spanish] *Economic Analysis Review* **29**(1): 57-72.

41. Medel, C.A. and P. Pincheira, 2015, "The Out-of-sample Performance of an Exact Median-Unbiased Estimator for the Near-Unity AR(1) Model," *Applied Economics Letters* **23**(2): 126-131.

42. Mise, E., T.-H. Kim, and P. Newbold, 2005, "On Suboptimality of the Hodrick-Prescott Filter at Time Series Endpoints," *Journal of Macroeconomics* **27**(1): 53-67.

43. Nason, J.M. and G.W. Smith, 2008, "Identifying the New Keynesian Phillips Curve," *Journal of Applied Econometrics* **23**(5): 525-251.

44. Newey, W.K. and K.D. West, 1987, "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* **55**(3): 703-708.

45. Orphanides, A., 2001, "Monetary Policy Rules Based on Real-Time Data," *American Economic Review* **91**(4): 964-985.

46. 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): 415-434.

47. Paloviita, M., 2009, "Estimating Open Economy Phillips Curves for the Euro Area with Directly measured Expectations," *New Zealand Economic Papers* **43**(3): 233-254.

48. Pesaran, M.H., T. Schuermann, and S.M. Weiner, 2004, "Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model," *Journal of Business and Economic Statistics* **22**(2): 129-162.

49. Pincheira, P. and R. Alvarez, 2009, "Evaluation of Short Run Inflation Forecasts and Forecasters in Chile," *Money Affairs* **XXII**: 159-180.

50. Pincheira, P. and C.A. Medel, 2012, "*Forecasting Inflation with a Random Walk*," Working Paper 669, Central Bank of Chile.

51. Pincheira, P. and H. Rubio, 2015, "The Low Predictive Power of Simple Phillips Curves in Chile," *CEPAL Review* **116**: 171-195.

52. Posch, J. and F. Rumler, 2014, "Semi-Structural Forecasting of UK Inflation Based on the Hybrid New Keynesian Phillips Curve," *Journal of Forecasting* **34**(2): 145-162.

53. Rabanal, P. and J.F. Rubio, 2005, "Comparing New Keynesian Models of the Business Cycle: A Bayesian Approach," *Journal of Monetary Economics* **52**: 1151-1166.

54. Rudd, J. and K. Whelan, 2005, "New Tests of the New-Keynesian Phillips Curve," *Journal of Monetary Economics* **52**(6): 1167-1181.

55. Rudebusch, G.D. and L.E.O. Svensson, 1999, *Policy Rules for Inflation Targeting*, in J.B. Taylor (Ed.), *Monetary Policy Rules*, University of Chicago Press, US.

56. Sbordone, A.M., 2002, "Prices and Unit Labour Costs: A New Test of Price Stickiness," *Journal of Monetary Economics* **49**: 265-292.

57. 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): 1123-1175.

58. 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): 161-183.

59. Smets, F. and R. Wouters, 2007, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review* **97**(3): 586-606.

60. 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.

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

62. Stock, J.H. and M.W. Watson, 1999, "Forecasting Inflation," *Journal of Monetary Economics* **44**(2): 293-335.

63. Stock, J.H. and M.W. Watson, 2004, "Combination of Forecasts of Output Growth in a Seven-Country Data Set," *Journal of Forecasting* **23**: 405-430.

A Statistical benchmarks

A.1 Univariate stationary autoregression

Alongside the RW, stationary AR models complement the most traditional benchmarks used for forecasting inflation as well as many other macroeconomic time-series (Ghysels, Osborn, and Rodrígues, 2006). The fitted models often include an moving average component (following the Box and Jenkins, 1970, model selection view); and so I refer to the ARIMA(p,1,0) particular case for simplicity. This also is due to the high persistence exhibited by inflation series, whose dynamics is well described by an AR(1) with a near-unity coefficient (see Pincheira and Medel, 2012, for details).

The strategy used in this article simply consists of estimating equation (A1) across the different p integers using the estimation sample. In this case, using $p^{\max}$=s=12 (s=annual frequency of the series) yields:

$$\pi_t = \overline{\pi} + \sum_{i=1}^{p\in P} \phi_i \pi_{t-i} + \varepsilon_t, \tag{A1}$$

where $\{\overline{\pi},\{\phi_i\}_{i=1}^{p\in P},\sigma_\varepsilon^2\}$ are parameters to be estimated, $\varepsilon_t \sim iid\mathcal{N}(0,\sigma_\varepsilon^2)$, and P={1,...,12}. For each "p"-model, the Bayesian Information Criterion (BIC) is computed whereas the forecasting model is that with the smallest BIC score (reflecting the better adjustment to the true model given the sample size). The BIC is defined as $\mathsf{BIC}{=}{-}2\mathcal{L}+(1+p)\log(T)$, where $\mathcal{L}$ is the log-likelihood function, T the sample size, and $(1+p)$ is the number of coefficients of the model (accounting: one constant plus p AR coefficients).

Many articles analyse the appropriateness of information criteria for forecasting purposes. Among the most used are the BIC, Akaike IC (AIC), the Hannan-Quinn, and the Mallows Cp Criterion. However, at least these four are derived under the same Kullback and Leibler (1951) principle of *cross entropy*, delivering the same asymptotic results. The BIC produces more parsimonious (in-sample) results with intermediate sample size compared to the AIC. But, this is still not sufficient to ensure higher out-of-sample accuracy. Moreover, Medel (2015c) finds that the overfitting is hazardous for forecasting accuracy when the number of parameters of the model exceeds at least the annual frequency of the series, *i.e.* when $p > s$. Hence, for the sake of parsimony, AR with BIC is preferred.

The ϕ_i-coefficient(s) are estimated made with the OLS method. This is in full acknowledgement of the downward bias that OLS provides for $\widehat{\phi}_i$ (see Lovell, 2008). Hence, no available bias-correction estimation is used including those of Andrews (1993) among others. This option is left because, as shown in Pincheira and Medel (2012) and Medel and Pincheira (2015), among the competing models is the RW, which results in a superior alternative for near-unity series. As the RW is used as a numerary model to compare the RMSFE, it results in a demanding benchmark for the economics-based models.

A.2 The exponential smoothing forecast

The ES corresponds *per se* to a forecasting model. The version used in this article corresponds to the *single* ES, but there are more specifications available, such as the *double* ES and the Holt-Winters model (see Hyndman *et al.*, 2008). The prediction for h-steps ahead is the same independently of the horizon:

$$\pi_{t+h|t} = \psi\pi_{t-1} + (1-\psi)\pi_{t-1+h|t-1}, \tag{A2}$$

with $0 < \psi \leq 1$. Note that if ψ=1, the ES coincides with the RW model. The model has been also used for forecasting purposes with relative success for the same reasons of the RW.

A.3 The random walk model

The RW consists of the special AR(1) case where ϕ is not estimated and it is restricted to ϕ=1 instead. This restriction, although simple, entails several methodological as well as economic consequences. The most significant impact is that it turns inflation into a non-stationary variable theoretically without available statistical inference and divergent predictions over the forecasting horizons. Due to this non-stationarity, it sounds unlikely–at least theoretically–to have room for stabilisation policymaking, since past unpredictable shocks do not vanish in time. Note that this argument is raised because inflation exhibits a unit root; hence, with a CPI$\sim$I(2). For forecasting purposes, it does not comprise a major setback since the solution of over-differentiation does not necessarily jeopardise the accuracy (Dickey and Pantula, 1987).

The empirical evidence has been overwhelmingly in favour of the RW. This is due to the benefit of misspecification that more than offset the parameter uncertainty arisen from finite sample estimation. This article uses a *driftless* RW forecast, following the argument given in Pincheira and Medel (2012) and Medel and Pincheira (2015) that driftless RW-based forecast are unbiased.

B Data

This appendix statistically describes the dataset used in this article. There are two kinds of data: inflation time series and the output gap, which is constructed using the Industrial Production (IP) index. The source of actual headline inflation and the IP of all countries is the *OECD Database*, whereas for inflation expectations it is the monthly *Consensus Forecasts* (CF) report prepared by *Consensus Economics*. I also use the RER index in a robustness exercise (source: *International Finance Statistics*, *International Monetary Fund*). Table B1, presents a detailed summary of the sources, measurement units in their original versions, plus the descriptor of each variable.

Table B1: Variable description (*)

Variable	Country	Unity	Scale	Descriptor	Source
Consumer Price	BRA	Index	2010=100	Consumer Prices - All Items	OECD Database
Index	CHL	Index	2010=100	Consumer Prices - All Items	OECD Database
(transformed	CHI	Index	2010=100	Consumer Prices - All Items	OECD Database
to inflation	EUR	Index	2010=100	Harmonised CP (19 countries)	OECD Database
series)	JPN	Index	2010=100	Consumer Prices - All Items	OECD Database
	US	Index	2010=100	Consumer Prices - All Items	OECD Database
Inflation	BRA	Basis points	None	Avg. % chg. on prev. yr	Consensus Economics
Expectations	CHL	Basis points	None	Avg. % chg. on prev. yr	Consensus Economics
	CHI	Basis points	None	Avg. % chg. on prev. yr	Consensus Economics
	EUR	Basis points	None	Avg. % chg. on prev. yr	Consensus Economics
	JPN	Basis points	None	Avg. % chg. on prev. yr	Consensus Economics
	US	Basis points	None	Avg. % chg. on prev. yr	Consensus Economics
Industrial	BRA	Index	2010=100	Production of total industry sa	OECD Database
Production	CHL	Index	2010=100	Production of total industry sa	OECD Database
(used for	CHI	Index	2010=100	Production of total industry sa	OECD Database
the output	EUR	Index	2010=100	Production of total industry sa	OECD Database
gap variable)	JPN	Index	2010=100	Total retail trade (volume)	OECD Database
	US	Index	2010=100	Production of total industry sa	OECD Database

(*) "sa" stands for seasonally adjusted. Source: Author's elaboration.

The whole sample span runs from 2000.1 to 2014.12 (180 observations). For in-sample modelling diagnostic checking, the first six years of observations (2000.1-2005.12) are used, and the remaining part for evaluation purposes (108 observations; 2006.1-2014.12). As abovementioned, the predictive ability of all the models is analysed with a *shortened evaluation sample* (2006.1-2008.8, 32 observations) for an analysis on model's behaviour prior to the crisis.

B.1 Inflation data

Note that the commodity prices boom of 2006-7 and the financial crisis of 2008-9 are included in the evaluation sample, making the task of forecasting more demanding. This is explicitly considered in this article using the *shortened evaluation sample*. This has to be considered when comparing with previous studies using a sample with smoother series.

The descriptive statistics of the inflation series considering the six countries are presented in Table B2 for three samples. Actual inflation is transformed using the annual percentage change of the CPI. This is made to fit the specification used by the expectation series. CF survey is entirely reported for the same transformation (for inflation variable); even if CPI-basket re-definitions will be undertaken. The expectation series are also the limiting variable for the sample span, starting in 2000. Inflation and IP (the latter analysed in Appendix C) are available in a useful quality since 1960s (assuming a backward reconstruction for the Euro Zone). Notice that for the full sample, it is presented the Augmented Dickey-Fuller (ADF) testing for stationarity. According to the ADF test, the inflation series are stationary at 5% of confidence, except Japan CF which is at 10% of confidence.

As the OE-HNKPC makes use of a weighting scheme, this article uses those coming from the first principal component. These weights are obtained with the full sample, but do not change dramatically with the estimation sample, and are presented in the "FLoading" row of Table B2. This is worth mentioning since a reliable forecasting exercise has to make use of the information conditional on the period in which it is available. For robustness, the forecasting exercise was re-do with an equally-weighted scheme delivering similar results. The factor loading reported includes the estimation with all the countries. Nevertheless, for each country-level estimation the weights are re-scaled to add to unity with a zero for the currently analysed country.

From Table B2, it is easy to notice why Brazil, Chile, and China concentrate close to 81% of the total variance of the inflation factor set. Particularly for the case of Brazil, the most of volatility is found in the estimation sample. Interestingly, and except for the case of the Euro Zone, both the mean and the variance of the series have increased during the evaluation sample–due to the two aforementioned episodes–also making the forecasting task more demanding. Another remarkable feature is that Japan exhibits a negative mean (and median) for the estimation sample, with a particularly low variance. Indeed, the behaviour of the Japanese CPI already corresponds to a stationary series.[10]

[10] However, as stated by Dickey and Pantula (1987), overdifferencing of the series does not carry an important issue when forecasting. In contrast, it is not recommended when the aim is to empirically test an economic theory.

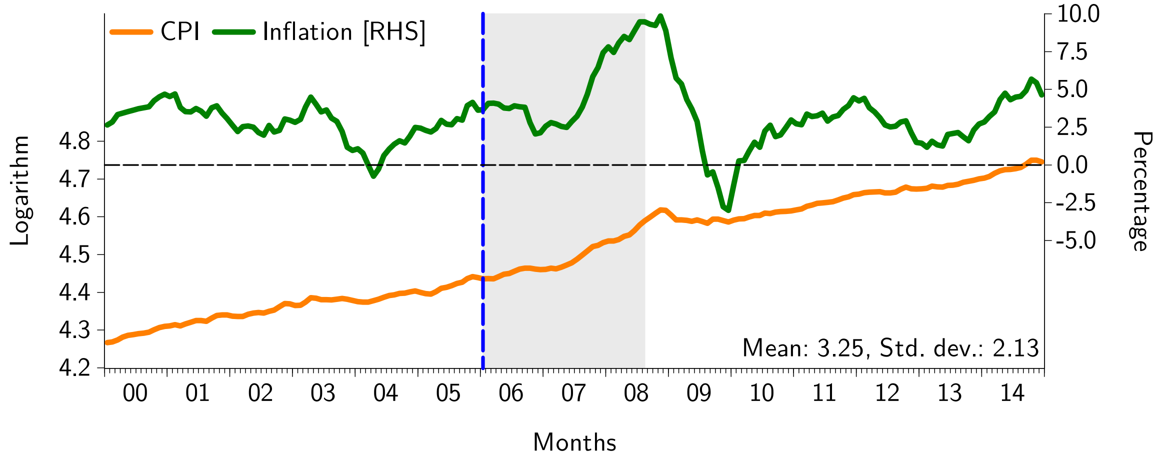
Table B2: Descriptive statistics of actual inflation series (*)

	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*	*BRA*	*CHL*	*CHI*	*EUR*	*JPN*	*US*
	Actual Inflation (π_t)						Consensus Forecasts (weighted) ($\widetilde{\pi}_t$)					
Full sample: 2000.1-2014.12 (180 observations)												
Mean	6.571	3.247	2.316	1.974	-0.032	2.379	5.542	3.178	2.549	1.561	0.077	2.094
Median	6.230	3.026	1.916	2.094	-0.200	2.317	5.300	3.096	2.704	1.563	-0.049	2.137
Max.	17.231	9.854	8.801	4.084	3.707	5.600	11.50	5.733	5.750	2.421	2.334	3.431
Min.	2.963	-3.011	-1.840	-0.645	-2.524	-2.097	3.600	1.550	0.150	0.307	-1.068	-0.448
Std. dev.	2.662	2.128	2.273	0.842	1.087	1.286	1.457	0.641	1.242	0.411	0.817	0.618
Skewness	2.261	0.706	0.608	-0.672	1.278	-0.585	2.356	0.997	0.046	-0.400	1.219	-1.291
Kurtosis	8.854	5.055	3.066	3.850	5.462	4.162	9.810	6.325	2.549	3.024	4.118	7.056
JB-Stat.	410.4	46.60	11.11	18.97	94.43	20.38	514.3	112.7	1.592	4.809	53.98	173.4
p-value	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.451	0.090	0.000	0.000
ADF-Stat.	-3.342	-3.658	-3.142	-3.301	-3.014	-3.700	-3.516	-4.279	-2.896	-3.257	-2.607	-3.771
p-value	0.014	0.005	0.025	0.016	0.035	0.004	0.008	0.007	0.047	0.018	0.093	0.003
FLoading	34.20%	27.14%	20.40%	12.94%	3.05%	2.27%	-	-	-	-	-	-
Estimation sample: 2000.1-2005.12 (72 observations)												
Mean	8.430	2.790	1.181	2.183	-0.480	2.689	6.343	3.076	1.756	1.398	-0.349	2.213
Median	7.396	2.953	0.984	2.136	-0.441	2.771	5.888	3.021	1.508	1.453	-0.294	2.188
Max.	17.213	4.694	5.250	3.123	0.797	4.687	11.50	4.046	3.750	1.822	0.228	3.061
Min.	5.155	-0.747	-1.217	1.609	-1.567	1.067	4.217	2.075	0.150	0.307	-1.068	1.570
Std. dev.	3.196	1.173	1.627	0.277	0.436	0.817	1.864	0.476	1.031	0.244	0.322	0.320
Skewness	1.674	-0.819	0.850	0.622	0.080	-0.126	1.709	0.111	0.362	-1.594	-0.206	0.442
Kurtosis	4.443	3.534	3.114	3.802	3.287	2.320	4.991	2.320	1.995	7.370	2.003	3.467
JB-Stat.	39.87	8.905	8.715	6.574	0.323	1.577	46.95	1.536	4.600	87.78	3.490	3.003
p-value	0.000	0.012	0.013	0.037	0.851	0.455	0.000	0.464	0.100	0.000	0.175	0.223
Evaluation sample: 2006.1-2014.12 (108 observations)												
Mean	5.332	3.552	3.073	1.834	0.267	2.173	5.008	3.246	3.077	1.670	0.361	2.014
Median	5.533	3.212	2.576	1.900	0.000	2.042	5.104	3.100	3.108	1.761	0.250	2.002
Max.	7.309	9.854	8.801	4.084	3.707	5.600	6.317	5.733	5.750	2.421	2.334	3.431
Min.	2.963	-3.011	-1.840	-0.645	-2.524	-2.097	3.600	1.550	0.467	0.581	-1.023	-0.448
Std. dev.	1.095	2.535	2.333	1.043	1.275	1.489	0.725	0.726	1.082	0.462	0.918	0.745
Skewness	-0.467	0.470	0.290	-0.230	0.717	-0.328	-0.079	0.973	-0.085	-0.813	0.662	-0.998
Kurtosis	2.323	3.723	3.000	2.490	3.742	3.409	1.962	5.760	3.465	2.866	2.744	5.001
JB-Stat.	5.993	6.336	1.514	2.120	11.73	2.693	4.958	51.32	1.101	11.97	8.184	35.94
p-value	0.050	0.042	0.469	0.347	0.003	0.260	0.084	0.000	0.577	0.003	0.017	0.000

(*) "JB-Stat." stands for Jarque-Bera test statistic (NH: Data are random). "ADF-Stat." stands for Augmented Dickey-Fuller test statistic (NH: Series has a unit root). ADF equations for π_t includes a constant with 4 lags (BRA, CHL, CHI, US), or 10 lags (EUR, JPN). ADF equations for $\widetilde{\pi}_t$ includes a constant with 4 lags (BRA, CHL, CHI, EUR, US) or 7 lags (JPN). Source: Author's elaboration.

Figure B1 plots both the CPI log-level and the annual percentage change for Chile; the target forecast variable. A quite different dynamic between the estimation and evaluation sample is easy to notice. While the mean achieves a lower 2.8% in the first part of the sample, the second increases to 3.5% (close to the inflation target), peaking at 9.8% in November 2008 and throughing at -3.0% in December 2009. The remaining inflation series are depicted in Figure B2, providing three salient features. The most obvious is the different dynamics in the CPI level of Japan, which is already stationary. A second feature is that for China, the Euro Zone, Japan, and the US a V-shaped pattern is observed in the inflation series during the 2008-9 period, which is the major episode contributing to the variance of the series.

Figure B1: Chilean Consumer Price Index. Log-level and annual percentage change (*)
Full sample



(*) Vertical line = evaluation sample start point. Shaded area = shortened evaluation sample.
Source: Author's elaboration.

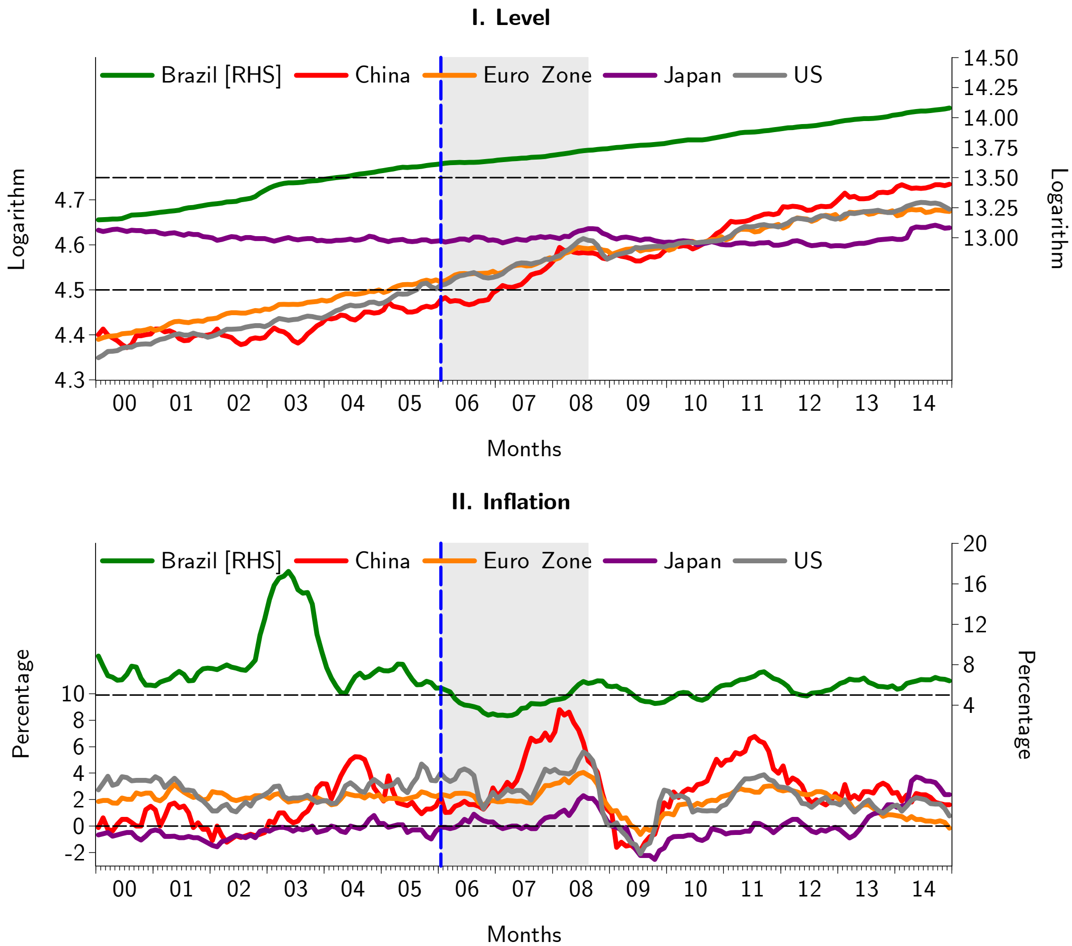
B.2 Inflation expectations data

The CF expectations are reported monthly, providing the point forecast of 15-20 agencies and private consultants for several variables at two fixed horizons: December of the current and the next year. The names of the respondents are explicitly revealed along with their forecasts, making possible a one-by-one accuracy analysis. Given this specific richness of the survey, several articles make use of CF for testing economic/statistic hypothesis. Interestingly, Pincheira and Alvarez (2009) jointly compare Chilean inflation forecasts reported by *Consensus Economics*, time series models, and those generated by Central Bank of Chile's staff.

However, as the estimation is made with constant frequency using recursive estimation, there is the need to adjust the series to have a unique rolling-event forecast. The approach used in this article is to create one series with a weighting scheme of the two forecasts in order to better accommodate the information to the targeted rolling-horizon. Hence, the CF forecast series for each month are weighted according to:

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Current Dec ($\widetilde{\pi}_t^{Current}$):	92%	83%	75%	67%	58%	50%	42%	33%	25%	17%	8%	0%
Next Dec ($\widetilde{\pi}_t^{Next}$):	8%	17%	25%	33%	42%	50%	58%	67%	75%	83%	92%	100%

Figure B2: Consumer Price Index time series. Log-level and annual percentage change (*)
Full sample



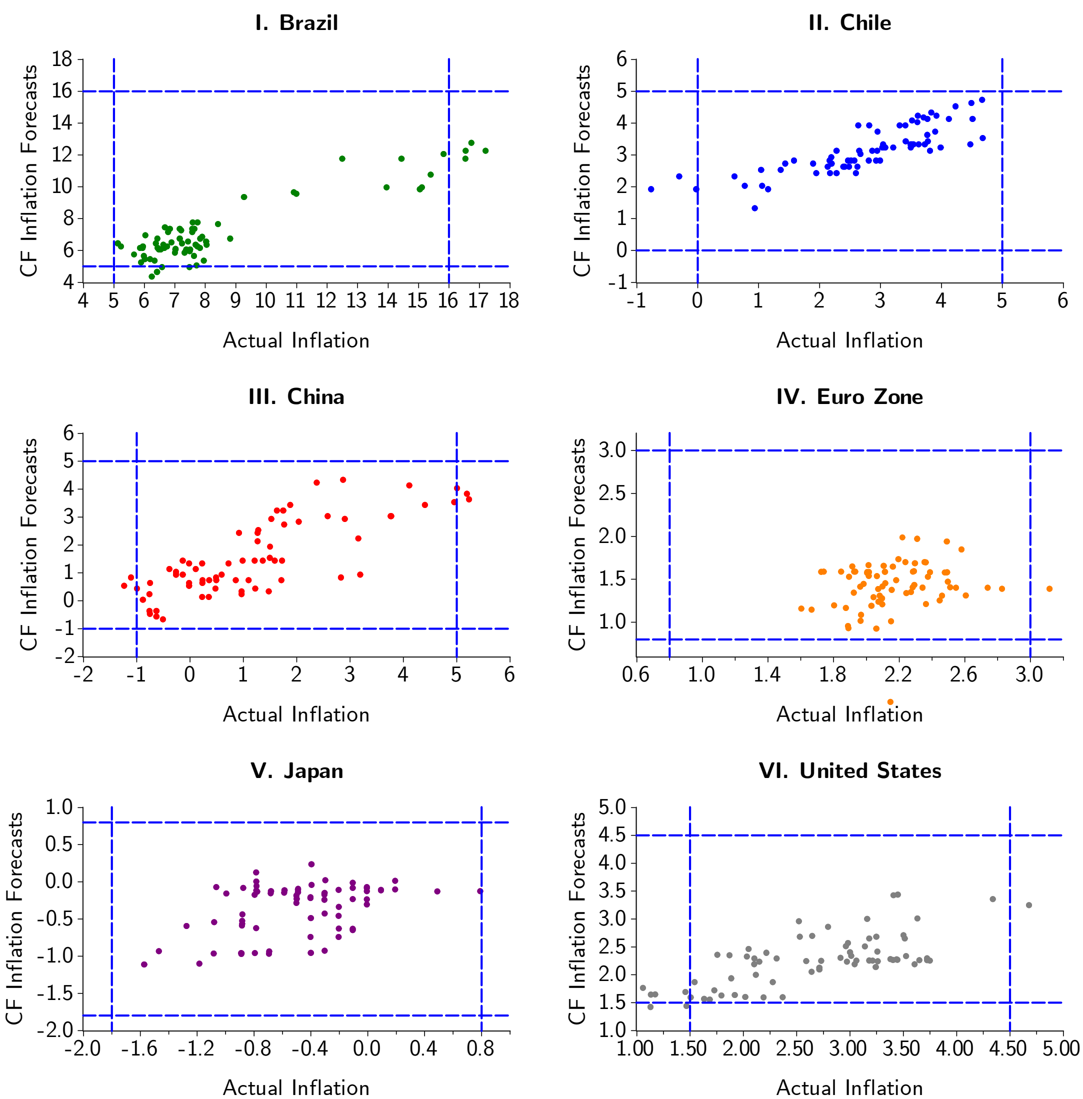
(*) Vertical line = evaluation sample starting point. Shaded area = shortened evaluation sample. Source: Author's elaboration.

Figure B3 presents the scatter plot between actual inflation and the CF for December of the current year for all the countries. The result, despite that CF is already accurate for the fixed-horizon, is no longer useful in a rolling-event scheme because the majority of the observations lie outside the 45° line. The Chinese CF series is the case that matches best the fixed horizon forecast with the rolling-event evaluation. However, this fact obeys just to a particular case, reinforcing the need to combine both expectations series into a unique measure. For the Chilean case it is found that the CF expectation for December of the current year consistently overestimates the inflation rate expected for the next 12 months when actual inflation is below 3% (the inflation target). But when the actual inflation is in the vicinity of the target, the expected inflation for December of the current year is close to that forecast 12-months-ahead. This fact, added to the results found in Medel (2015b)–that Chilean SPF's expected inflation 24 months ahead is consistently equal to the target–can be read as strong confidence of the forecasters to the commitment of the central bank to its mandate.

The last six columns of Table B2 show the descriptive statistics of the weighted CF series. In this case, and judging by point estimates (mean and median) the accuracy is notably improved across the sample.

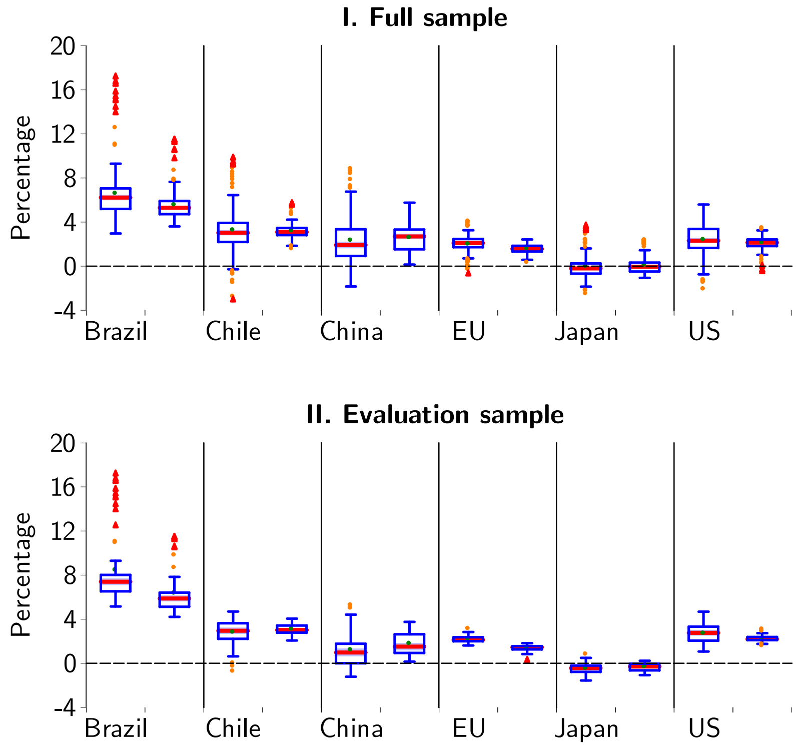
A more suitable way to visualise this is presented in the boxplots of Figure B4. In Figure B4 there are ix pairs of boxplots, each pair showing first the actual and then the CF (weighted) statistics using the full and estimation sample. The most salient feature is the reduced number of outliers in the evaluation sample. Note that the CF weighted series fulfils three desirable features in a forecast series: the mean (green dot) is close to the mean of the actual series, the volatility (proxied with the width of the blue box) is smaller than that of the actual series, and finally, CF exhibits fewer outliers (orange and red dots) than the target variable. This last feature is particularly easy to notice with the evaluation sample.

Figure B3: Scatter plot of CF inflation forecasts for December of current year (*)
Evaluation sample



(*) Source: Author's elaboration.

Figure B4: Descriptive statistics of actual inflation and weighted inflation forecasts (*)



(*) Source: Author's elaboration.

C Output gap building blocks

This appendix follows closely the output gap construction used in Medel (2015a, 2015b). One of the major drawbacks when estimating the NKPC is the impossibility to accurately measure the excess demand–*i.e.* marginal costs. As the CE-HNKPC and the OE-HNKPC make use of this measure, it is desirable to have a stable series as new observations are added. The typical alternative to the marginal cost variable is the output gap ($\widetilde{y}_t$)–*i.e.* the difference between current and potential output.[11] As the estimations are made with monthly data, the IP index is used as a proxy of the quarterly GDP. Table C1 presents the descriptive statistics of these series for all countries and for the two sample spans, for the annual percentage change (Δ^{12}) of the level series.

Note that the transformation achieves stationarity according to the ADF test. The statistics of Table C1 remarkably describe the textbook result on growth convergence. In other words, industrialised countries grow less than developed ones because the former are closer to a steady state than the latter. The Euro Zone, Japan, and the US exhibit an average rate of growth not greater than 1.5%, whereas for Brazil and Chile this rate achieves 4.5 and 3.8%, respectively. For China, the average rate achieves an astonishing 14%, with a standard deviation similar to that calculated for Brazil and Chile. Graphically

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

(not shown) all the series exhibit the same V-shaped behaviour during the financial crisis, coinciding with the maximum and minimum value reported in Table C1.

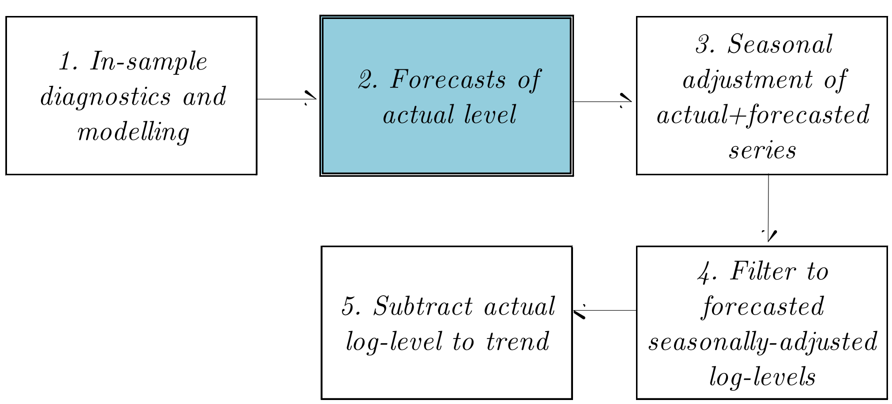
Basically, instability in the output gap arises with the "end-of-sample" problem of filtering, especially when the Hodrick-Prescott (HP) procedure is used to obtain the potential output: an unobservable component.[12] To alleviate this setback, I follow the approach proposed by Bobbitt and Otto (1990), Kaiser and Maravall (1999), and more recently re-launched by Mise, Kim, and Newbold (2005). This consists of adding forecast observations to level series prior to performing any filtering procedure. Hence, the method applied to obtain the output gap follows the steps of Figure C1 where the enhancements start in the second shaded block. Note that the seasonal adjustment is made with X12-ARIMA in its default mode, and the filtering method is HP (λ=129,600).

Table C1: Descriptive statistics of Industrial Production time series (*)

	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*	*BRA*	*CHL*	*CHI*	*EUR*	*JPN*	*US*
	Industrial Production (y_t)											
	Estimation sample						*Evaluation sample*					
Mean	3.775	4.545	13.99	1.483	1.389	1.331	1.416	1.571	12.86	0.015	-0.174	1.101
Median	4.120	4.624	14.40	1.237	2.736	2.065	2.048	1.981	12.80	1.450	1.905	2.708
Max.	11.282	14.61	23.20	7.302	7.794	5.429	19.69	30.91	20.10	9.296	27.32	8.519
Min.	-6.476	-4.367	2.300	-3.832	-12.76	-5.811	-16.45	-17.49	5.400	-21.62	-33.33	-15.42
Std. dev.	3.855	3.976	3.958	2.455	5.267	2.908	6.716	5.997	3.665	6.666	10.73	5.157
Skewness	-0.155	-0.031	-0.486	0.159	-1.260	-0.930	-0.036	0.674	0.165	-1.443	-0.718	-1.874
Kurtosis	2.598	2.981	3.426	2.661	3.661	3.000	3.970	8.097	1.960	4.888	4.812	5.724
JB-Stat.	0.773	0.012	3.380	0.650	20.37	10.377	4.256	125.1	5.350	53.49	24.052	96.61
p-value	0.679	0.994	0.185	0.722	0.000	0.006	0.119	0.000	0.069	0.000	0.000	0.000
ADF-Stat.	-3.845	-3.213	-3.618	-4.228	-4.121	-3.083	-	-	-	-	-	-
p-value	0.003	0.020	0.006	0.000	0.001	0.029	-	-	-	-	-	-

(*) "JB-Stat." stands for Jarque-Bera test statistic (NH: Data are random). "ADF-Stat." stands for Augmented Dickey-Fuller test statistic (NH: Series has a unit root). ADF equation includes a constant and 1 lag (CHI), or 4 lags (BRA, CHL, EUR, JPN, US), using the full sample. Source: Author's elaboration.

Figure C1: Output gap building blocks (*)



(*) Source: Author's elaboration.

[12] See Orphanides (2001) for a discussion on this matter.

The ARMA forecasting model for IP corresponds to $\Delta^{12}y_t = \overline{y} + \phi\Delta^{12}y_{t-p} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_{12}\varepsilon_{t-12} + \theta_1\theta_{12}\varepsilon_{t-13}$, with $\varepsilon_t \sim iid\mathcal{N}(0,\sigma_\varepsilon^2)$. This is a version of the so-called *airline model* (Box and Jenkins, 1970) which has proved to be a model that fits macroeconomic data with substantial success (Ghysels, Osborn, and Rodrígues, 2006). The in-sample estimates are presented in Table C2, which also reveals robust results across countries, and a correct specification according to the Durbin-Watson statistic, defined as $\mathsf{DW}{=}\Sigma_{t=2}^{T}(\varepsilon_t - \varepsilon_{t-1})^2/\Sigma_{t=1}^{T}\varepsilon_t^2 \approx 2(1-\rho_\varepsilon)$, where ρ_ε is errors' autocorrelation.

Several articles use the output gap as a proxy of marginal costs, differing often on the way to obtain de-trended output (whether based on HP or other filtering device). The economic rationale behind this measure is striking; it considers the distance between the current state of the economy and the counterfactual that may be obtained if all factors were employed in the absence of shocks. Some examples using the output gap are Rudebusch and Svensson (1999), Stock and Watson (1999), Galí, Gertler, and López-Salido (2005), Lindé (2005), Paloviita and Mayes (2005), Rudd and Whelan (2005), Canova (2007), Dees *et al.* (2009), and Jean-Baptiste (2012), among others. Moreover, Batini, Jackson, and Nickell (2005) use the output gap alongside the labour share on the basis of an endogenously determined price mark-up.

Table C2: In-sample diagnostics of IP forecasting models (*)

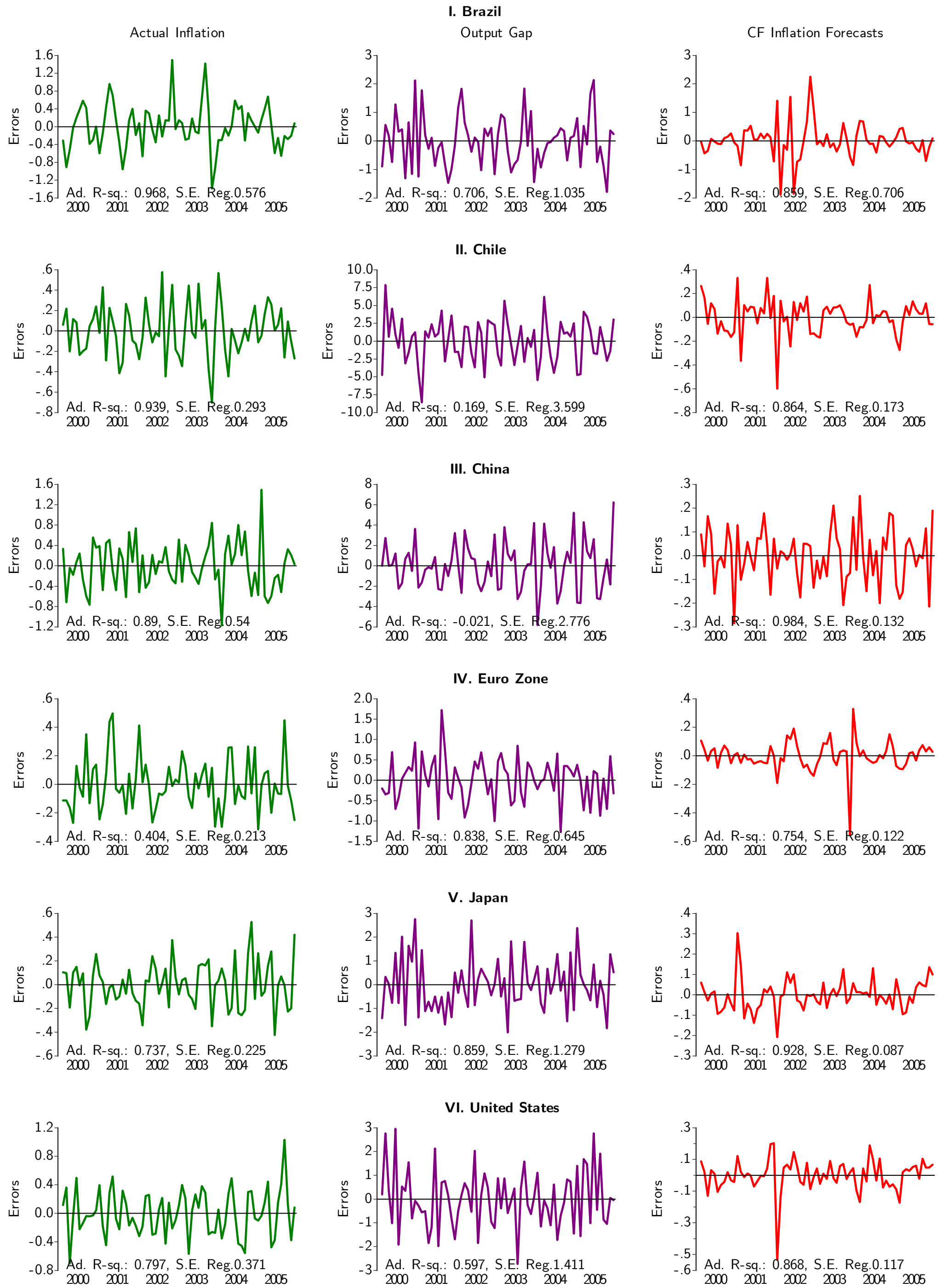
	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*
	Dependent variable: $\Delta^{12}y_t$					
	Estimation sample					
ϕ	-0.276	-0.790	-0.671	-0.336	-0.381	-0.445
	[0.031]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
θ_1	-0.896	-	-	-0.898	-0.900	-
	[0.000]	-	-	[0.000]	[0.000]	-
θ_2	-	-0.752	-0.633	-0.898	-0.900	-
	-	[0.000]	[0.000]	[0.000]	[0.000]	-
θ_3	-	-	-	-0.898	-0.900	0.221
	-	-	-	[0.000]	[0.000]	[0.007]
θ_{12}	0.527	0.882	0.833	0.619	0.455	0.889
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$\overline{y}$	0.204	0.301	0.400	0.137	0.101	0.149
	[0.001]	[0.000]	[0.000]	[0.000]	[0.066]	[0.490]
$\overline{R}^2$	0.229	0.778	0.759	0.216	0.286	0.701
S.E. Reg.	0.962	2.711	1.721	0.759	1.075	1.179
DW Stat.	2.012	2.098	2.181	2.208	1.731	1.994

(*) Equation: $\Delta^{12}y_t = \overline{y} + \phi\Delta^{12}y_{t-p} + \varepsilon_t + \theta_p\varepsilon_{t-p} + \theta_{12}\varepsilon_{t-12} + \theta_p\theta_{12}\varepsilon_{t-12-p}$ with $\varepsilon_t \sim iid\mathcal{N}(0,\sigma_\varepsilon^2)$. See notes to Table 1.
Source: Author's elaboration.

Stock and Watson (1999) suggest that, especially when the aim is to forecast, the output gap measure provides a convenient alternative since it relies basically on a univariate ensemble. Also, some of the major problems associated with output gap–instead of using marginal cost–are rather an empirical issue. The forecasts provided by the models of Table C2 tackle part of the "end-of-sample" problem.

D GVAR residuals diagnostics

Figure D1: GVAR residuals time series (*)
Estimation sample

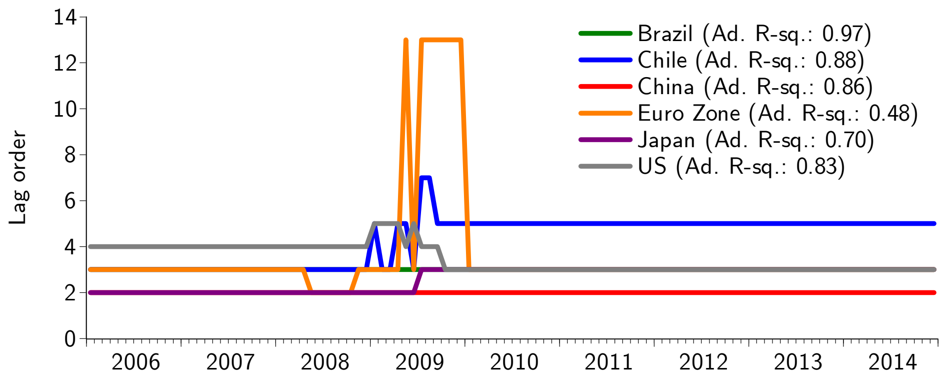


(*) Source: Author's elaboration.

E Autoregression in-sample diagnostics

Figure E1 presents the chosen lag profile across the time of the AR models (showing always p lags plus one constant). This profile arise since a rolling-window sample scheme is used. The lag length is chosen according to the BIC. As expected, and in line with the CE-HNKPC estimates, during the 2008-9 more coefficients are required by the model to capture the volatile behaviour of the series. During this period, the Euro Zone achieves the maximum number of lags allowed. In Table E1, the first point estimation of all AR models are presented showing significant coefficients and that they are well specified according to the DW statistic.

Figure E1: AR chosen lag length profile across evaluation sample (*)
Evaluation sample



(*) Source: Author's elaboration.

Table E1: AR models' diagnostics (*)

	BRA	*CHL*	*CHI*	*EUR*	*JPN*	*US*
	Dependent variable: π_t					
	Estimation sample					
ϕ_1	1.684	1.235	0.922	0.858	0.841	1.188
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
ϕ_2	-0.737	-0.321	-	-0.286	-	-0.664
	[0.000]	[0.003]	-	[0.034]	-	[0.000]
ϕ_3	-	-	-	-	-	0.401
	-	-	-	-	-	[0.000]
$\overline{\pi}$	8.399	2.832	1.473	2.192	-0.456	2.717
	[0.000]	[0.000]	[0.170]	[0.000]	[0.006]	[0.000]
$\overline{R}^2$	0.971	0.881	0.855	0.477	0.704	0.827
S.E. Reg.	0.548	0.408	0.619	0.199	0.237	0.341
DW Stat.	1.849	1.885	1.864	2.041	1.843	1.817

(*) Equation: $\pi_t = \overline{\pi} + \phi_1\pi_{t-1} + \phi_2\pi_{t-2} + \phi_3\pi_{t-3} + \varepsilon_t$, with $\varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$. See notes to Table 1.
Source: Author's elaboration.