



## Searching for the Best Inflation Forecasters within an Employment Survey: Microdata Evidence from Chile

Carlos A. Medel<sup>a</sup>

<sup>a</sup>Central Bank of Chile

✉ cmedel@bcentral.cl

### Abstract

This article aims to evaluate quantitative inflation forecasts for the Chilean economy, taking advantage of a specific survey of consumer perceptions at the individual microdata level, which, at the same time, is linked to a survey of employment in Chile's capital city. Thus, it is possible to link, with no error, consumer perceptions and 12-month-ahead inflation forecasts with personal characteristics such as gender, age, educational level, county of living, and the economic sector in which they are currently working. By using a sample ranging from 2005.II to 2018.IV, the results suggest that women aged between 35 and 65 years old, with a college degree, living in the North-eastern part of Santiago (the richest of the city), and working in the *Community and Social Services* sector are the best forecasters. Men aged between 35 and 65 years old, with a college degree, in a tie living in the South-eastern and North-eastern part of the city but working in *Retail* and *Government and Financial Services* sectors, respectively, come in second. Some econometric exercises reinforce and give greater support to the group of most accurate forecasters and reveal that another group of forecasters, different from the second-best in terms of forecast accuracy, displays the characteristics required of a forecasting variable. Remarkably, this group has the same specifications as the most accurate group, with the only difference being that it is composed of men instead of women. Thus, it looks promising for further consideration. Importantly, a forecast accuracy test reveals that no factor comes out as superior to the naïve random walk forecast used as a benchmark. These results are important because they help to identify the most accurate group when forecasting inflation and, thus, help refine the information provided by the survey for inflation forecasting purposes.

**Article History:** Received: 13 May 2021 / Revised: 07 September 2021 / Accepted: 03 December 2021

**Keywords:** Employment survey; Inflation; Consumer sentiment; Microdata; Forecasting; Survey data

**JEL Classification:** C53, C83, E31, E37

## Acknowledgements

I thank the comments and suggestions of Pilar Cruz, Carla Fucito, Carlos Madeira, Pablo Medel, two anonymous referees, Guest Editors, participants at the Internal Central Bank of Chile's Workshop on "Unstructured Data Analysis for Macroeconomics and Monetary Policy", held in September 2021, Fernando Faure for superb research assistance, Cynthia van Lamoen from Universidad de Chile's *Centro de Microdatos*, and Consuelo Edwards for editing services. I also thank Mario Marcel—to whom this paper is dedicated—for drawing my attention to the database explored in this study. The views and ideas expressed in this article do not necessarily represent those of the Central Bank of Chile or its authorities. Any remaining errors or omissions are the author's sole responsibility.

## 1. Motivation, Objective, and Main Results

Given the lag with which monetary policy typically operates on the economy, accurate forecasts of the future state of the economy, and future inflation outcomes, are key for the success of monetary policy in particular and economic policy in general. Inflation expectations are, thus, one of the key pillars of the current understanding of monetary policy (Clarida, 2012; Mishkin, 2014, Chapter 13). Moreover, for central banks with a clear mandate on keeping inflation controlled (i.e., low, stable, and predictable) it is key to screen inflation expectations of different agents of the economy, such as analysts, observers, investors, regulators, consumers, and households. Furthermore, central banks operating within an expected inflation targeting scheme with firmly anchored inflation expectation at the policy horizon (e.g., 2-year as in the case of Chile), intermediate horizons such as 12-month ahead, become more important for the diagnosis of the economy as well as for the monetary policy stance.<sup>1</sup>

Similarly, it is well documented that consumer perceptions surveys provide rich, timely, and useful information from a representative group of agents of the economy that, more than often, require some statistical refinement to disentangle its purely forward-looking content. On many occasions, consumer perceptions surveys, aiming to capture the consumers' mood or sentiment about the economy, also include quantitative indicators such as the prospective inflation rate of the Consumer Price Index (CPI).<sup>2</sup>

This article aims to evaluate quantitative inflation forecasts for the Chilean economy, taking advantage of a specific survey of consumer perceptions at individual microdata level, which, at the same time, is linked to a survey of employment and unemployment of Chile's capital city, Santiago. Thus, the key advantage of the database is that it is possible to link, with no error, consumer perceptions and 12-month-ahead inflation forecasts<sup>3</sup> with labour-market characteristics

<sup>1</sup>Recent studies analysing consumers' inflation expectations outcomes as informational input for monetary policymaking are Coibion et al. (2020), Goldfayn-Frank et al. (2020), and Duca-Radu et al. (2021).

<sup>2</sup>See, for instance, the *Business and Consumer Survey* by the European Commission, *Eurobarometer*, the *Survey of Consumers* by the University of Michigan, the New York Federal Reserve Bank's *Survey of Consumer Expectations*, the *Canadian Survey of Consumer Expectations*, the *Swiss Consumer Sentiment* survey by the State Secretariat for Economic Affairs, and the *Opinion Survey* by the Bank of Japan for advanced economies as well as the Brazilian IBRE/FGV *Consumer Confidence Survey* and the Indian *Inflation Expectations Survey of Households* (IESH) in emerging market economies.

<sup>3</sup>In particular, I use the results of Question 8: "What will be the CPI inflation rate in 12 months?" (numeric

of the respondents, which heavily relies on personal (yet anonymous) characteristics.<sup>4</sup> This is done by constructing sub-sets of inflation expectations factors with fully identifiable and mutually excluded characteristics. Mentioned characteristics are gender, age, education, county of living, and the economic sector of present work, representing the exploitable personal attributes of the database. There are some other interesting attributes available, but with very small sample size, not suitable to conduct a reliable statistical inference.

The question of which group of consumers is better at inflation forecasting is not new.<sup>5</sup> However, there is scarce literature for emerging economies, particularly in the case of Chile. In this sense, this study constitutes the first analysis of Chilean inflation forecasts considering the respondents' personal, demographic, and geographic characteristics in an out-of-sample evaluation. However, few studies analyse Chilean inflation expectations inferred from both market-based financial securities and surveys, especially aiming to analyse the degree of expectations anchoring. Despite using a very short sample span (2002–2005) from a then non-developed Chilean financial market, Gürkaynak et al. (2007) provide supporting evidence that inflation breakeven obtained from financial assets remain anchored to the recently implemented inflation targeting regime. Larraín (2007), in turn, finds that it is not a clear cut how well inflation breakeven comes out as a good inflation expectations measure using the same sample span. Using two financial models, the author finds that mentioned market-based measure includes other components, such as risk and term premia, that could move in opposite directions to inflation expectations. However, with a larger sample span from a more developed Chilean financial market, De Pooter et al. (2014) use survey-, market-, and inflation-linked bonds to analyse the anchoring degree in Brazil, Chile, and Mexico. The authors find that inflation expectations have become much better anchored from 2004 onwards in the three economies. Particularly in the case of Chile, inflation expectations come out as less sensitive to economic news in China and the United States, which is understood as a strong anchoring degree. On a similar verge, Medel (2018) finds that inflation expectations from the *Economic Expectations Survey* conducted by the Central Bank of Chile display a low sensitivity to actual inflation readings based on a battery of econometric estimates. Using a dynamic stochastic general equilibrium (DSGE) model allowing for a time-varying learning mechanism of accumulated inflation forecast errors, Arias and Kirchner (2019) find that long-term expectations are insensitive to the arrival of new information in an episodic

---

answer). Despite that, the question is asked freely (and this is how they are found in the original *Stata* files of the database), in the sense that respondents are not forced to frame their answers in a probability distribution of possible outcomes, this is how the results are presented in the *Survey of Perception and Expectations on the Economic Situation in Greater Santiago* report in tranches of (-,2%), (2%,3%), (3%,4%), and (4%,+).

<sup>4</sup>Naturally, more surveys are asking for inflation expectations at different horizons to different agents in Chile. See Appendix A for a comparison of the expectations analysed in this study and others of common use.

<sup>5</sup>See, for instance, the linkages between demographic characteristics and inflation expectations in Bryan and Venkatu (2001a,b), Souleles (2004), Pfajfar and Santoro (2008), Coibion and Gorodnichenko (2015), Madeira and Zafar (2015), Ehrmann et al. (2017), Armantier et al. (2017), Axelrod et al. (2018), and Lyziak and Sheng (2018) for the case of the United States, Lindén (2005) and Drakos et al. (2020) for a group of European countries, Campelo et al. (2019) for the case of Brazil, Gosselin and Khan (2015) for Canada, Sabrowski (2008) for Germany, Goyal and Parab (2019a) for India, Malgarini (2009) and Easaw et al. (2013) for Italy, Ichiue and Nishiguchi (2015) and Diamond et al. (2020) for Japan, Rossouw et al. (2011) for South Africa, and Jonung (1981) and Palmqvist and Strömberg (2004) for Sweden.

manner. This is the probability that inflation expectations may become unanchored varies with accumulated inflation surprises, which is consistent with the learning mechanism in the proposed DSGE model. Finally, a recent review of Chilean inflation dynamics, including its expectations anchoring degree, is presented in the [Central Bank of Chile \(2020\)](#), especially considering the benefits provided by the inflation targeting regime adopted in 2000 in the form of a fixed target of 3% over a 2-year horizon.

By using a quarterly sample ranging from 2005.II to 2018.IV (55 observations), and different combinations of attributes leading to 648 different inflation forecasts (or “factors”), the results for total inflation suggest that women aged between 35 and 65 years old, with a college degree, living in the North-eastern part of the city (that with the highest living standards of the country), and working in *Community and Social Services* sector are the best forecasters of total inflation (i.e., the “winner factor”). Out of 3,060 consumers surveyed each quarter, this group comprises up to 26 consumers, i.e., the top 1% of forecasters. Men aged between 35 and 65 years old, with a college degree, in a tie living in the South-eastern and North-eastern part of the city but working in *Retail* and *Government and Financial Services* sectors, respectively, are the second-best at forecasting inflation. Finally, men aged between 35 and 65 years old, with a college degree, living in the North-western part of the city, working on *Government and Financial Services* are in the fourth place. From the fifth to eighth places, I found men aged between 35 and 65 years old as a common characteristic and concentrated in *Community and Social Services* and *Government and Financial Services*, but with a different spatial distribution. Notice that all these results are compared to the naïve random walk (RW) forecast, as it assumes no deeper knowledge on inflation dynamics, and thus, it serves as a predictive benchmark such as the cases of [Kenny et al. \(2014\)](#) using the University of Michigan’s *Survey of Consumers* for the United States, [Goyal and Parab \(2019b\)](#) using the *Consumer Confidence Survey* for India, and [Drakos et al. \(2020\)](#) for 17 European countries using *Eurobarometer*’s inflation forecasts.<sup>6</sup> Just the mentioned eight out of 648 total possible forecasts outperform the benchmark, with none of them coming out as statistically superior to the RW.

<sup>6</sup>Other reasons why the RW would be adequate as a benchmark. First, despite that inflation is stationary, its forecasts could be more accurate by using the RW model instead of, for instance, forecasts based on an estimated  $p$ -order autoregressive  $AR(p)$  model. This is because a highly persistent time series—such as the case of inflation—spoils out the estimation of near-unity coefficients (easily delivering suboptimal forecasts). Thus, the estimation bias is higher enough to more-than-offset model uncertainty with a finite sample. A misspecified model (such as the RW) performs systematically better than an estimated  $AR(p)$ . This is the key point of [Medel and Pincheira \(2016\)](#) and [Pincheira and Medel \(2016\)](#). Second, the RW is often referred to as “the naïve forecast” in the forecasting literature precisely because it is the type of forecast that does not require any knowledge of the underlying data generating process (or neither of econometric forecasting nor the current state of the macroeconomy). Instead, it is required (and sufficient to this end) to know the latest available observation only. This is a reasonable criterion for discriminating between informed and non-informed respondents about the macroeconomy without necessarily being experts in inflation forecasting. Third, the RW forecast is the same for all horizons, which is presumably the case for a non-expert sample of individuals that associate “12-month ahead” with a blurry “future horizon”, as mentioned in [Chanut et al. \(2019\)](#). Thus, this sort of “rational inattention” of the forecasting horizon is more likely to occur with a non-expert group of individuals; presumably, the case with the used dataset, or at least, is an adequate assumption when they asked to respond to a “consumer”. Finally, it is unclear what atheoretical model could be better used as benchmark than the RW, given the unknown knowledge on inflation by the respondents.

Also, several econometric exercises are conducted to further discriminate between the inflation factors that prove to be more accurate to the RW. These exercises (i.e., a brief comparison of key descriptive statistics, a regression-based biasedness analysis, forward- versus backward-lookingness estimates, a hybrid New Keynesian Phillips curve forecast comparison, and the U-Theil decomposition) reinforce and give more support to the winner inflation factor. However, these exercises reveal that a factor different from the second-best in terms of forecast accuracy displays the characteristics required of a forecasting variable. Remarkably, this mentioned factor has the same specifications as the winner factor, with the only difference being that it is composed of men instead of women. The results presented in this article are important because they help identify the most accurate group when forecasting inflation and, thus, help refine the information provided by the survey for inflation forecasting purposes.

The rest of the article describes the econometric setup, compounded by the dataset and the forecast evaluation framework. In Section 3, I present the results plus some econometric estimations to characterise and emphasise the differences between inflation forecasts, whereas in Section 4, I provide a discussion on three topics of interest in the light of results: (i) the extent to which the characteristics of the best forecasters are similar to that suggested by the international evidence, (ii) a discussion on why the mean is the chosen statistic instead of the median to represent individuals with the same characteristics, and (iii) an exploratory analysis on what type of price series the group with the less accurate results for the total inflation is targeting at. Finally, I conclude in Section 5.

## 2. Econometric Setup

### 2.1 Data

As mentioned above, I make use of the microdata freely available (after submitting an online registry) by the *Centro de Microdatos* of Universidad de Chile (<http://documentos.microdatos.cl/>). This database is the result of merging two surveys: the *Survey of Employment and Unemployment in Greater Santiago* (the EOD) and the *Survey of Perception and Expectations on the Economic Situation in Greater Santiago* (the IEE). A unique feature is that both databases are available (anonymously) at the individual level and are already merged. Respondents are asked about their labour situation as workers and their economic perceptions as consumers. The independence in answering both surveys, especially that of sentiment, is ensured by the wording of the questions. Thus, the sentiment is not conditioned to the labour situation by survey design.

Notice that the inflation expectation question is embedded in an employment survey, which typically applies the same questionnaire with a fixed frequency to a previously defined sampling scheme composed of individuals. On some occasions, like the one at hand, the unit of analysis are households and, therefore, representative expansion factors to match the total population is used. In the case of this survey, individuals are asked if they are responsible for a household—typically, the main earners—(“*Jefa/e de Hogar*”). If so, an expansion factor is applied according to their representativeness in the population. These expansion factors are representative according to the 2002 Census, as explained in *Centro de Microdatos* (2016). The corresponding next Census to

update the sample should be that of 2012. However, the reliability of its results—due to relevant methodological flaws—and a subsequent summarised version conducted in 2017 does not provide the necessary and reliable information to update the survey’s sampling scheme. Thus, I prefer to use the “whole sample” (i.e., that using expansion factors) as the data would be more likely used, as an employment survey, and as a lack of a better Census alternative. As an employment survey, it is more likely to be read (and used) measured in “number of individuals” (i.e., the labour force) to then calculate the different rates and flows of the labour market (e.g., unemployment rate).

Naturally, the universe of the IEE is the same as that of EOD:<sup>7</sup> inhabitants over 14 years old living in the Santiago Metropolitan Region and in Puente Alto and San Bernardo counties. This adds up to about 40% of the Chilean population in 2017. The sample comprises 3,060 individuals per quarter, consisting of stratified random sampling with a panel data component: a rotating panel. This is a method where part of the panel is kept permanently, and another part is an entirely new cross-section sample. The sample of 3,060 individuals in each quarter is divided into four subsamples of 765 individuals, where each subsample is independent and represents the Greater Santiago. The rotation design is of the 2-2-2 type, where two selected individuals are interviewed twice in a row. They are not contacted in the following two rounds, and then they are interviewed again in the following two rounds, covering a period of 18 months in total. However, the unique identificatory is not available to the public, and thus, it is only possible to conduct panel estimations considering a group of individuals with a common characteristic. The collection technique is face-to-face interviews, and the reported answer rate reaches 77.4% (informed in March 2014). The representativeness with respect to the universe could be considered adequate. It fulfils all sampling requirements, but this is not the case when the whole country is considered, as it focuses on Santiago only.<sup>8</sup> The IEE has been released quarterly and fully available since March 2001 (75 observations available until September 2019).

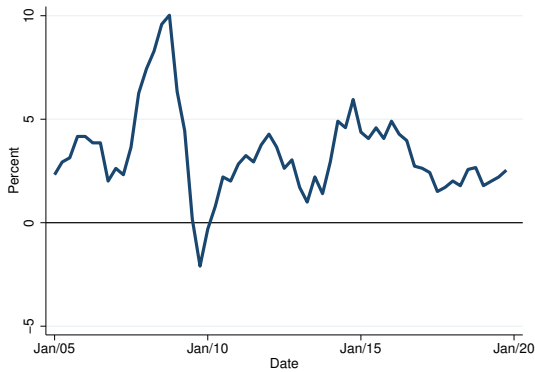
The merged database is compounded with a total of 142 variables. Out of this total, 18 are associated with the IEE, and 46 with the EOD. The remaining variables are answers on the household’s income and debt-related issues and working variables for internal use. However, financial variables are available in a shorter sample span, making it difficult to use them for this article. Moreover, not all series of the EOD are possible to use for this same reason or a very low answer rate. This is the case of income and the time of the respondent working in the same job. Particularly, these two questions would be useful to discriminate between groups (income as so, and the second as a proxy experience), and thus, other variables must fulfil this task.

The actual total inflation measure is presented in [Figure 1a](#). Given that inflation rates are presented in monthly frequency, two versions of the quarterly series are analysed: the end-of-period rate (comparing the annual variation of the last month of each quarter with the same month of the previous year), and the average rate (comparing the annual variation of the 3-month average of the quarter with the same average of the previous year). The variables used to classify and compound the inflation factors are gender, age, education, county of living, and economic

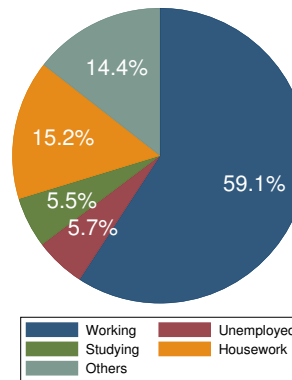
<sup>7</sup>Specific details can be found in *Centro de Microdatos* (2016).

<sup>8</sup>Santiago is Chile’s capital city with 7.037 million inhabitants of a country totalling of 18.730 million in 2017, thus representing 37.57% of the total population.

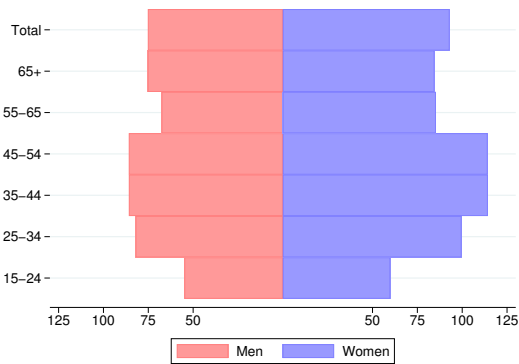
sector of current work. Notice that all these attributes are mutually exclusive. I consider all individuals that respond “Working” to the “Employment Situation” question because it leads to



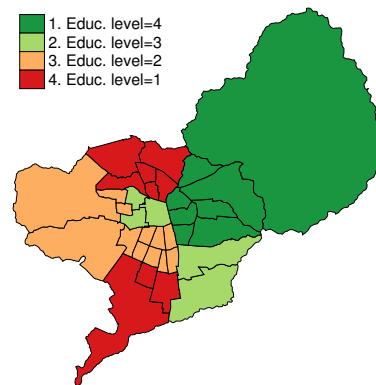
(a) Total inflation (annual variation).



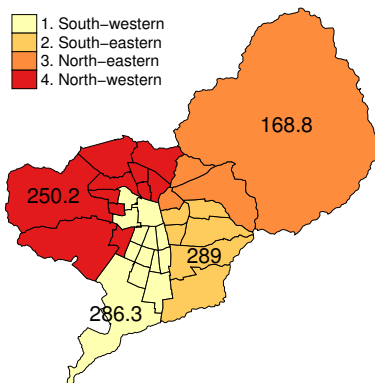
(b) Employment situation (percentage of total answers, 2005–18).



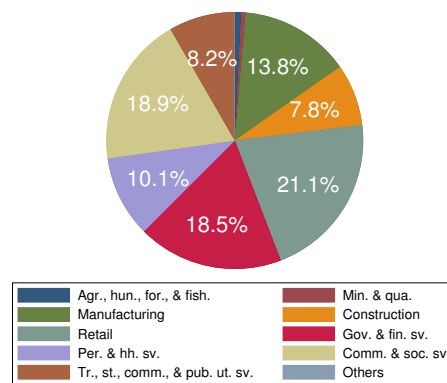
(c) (Original) age groups by gender (number of respondents, 2005–18 average).



(d) Education by (original) zones (8 zones) (re-coded education, 2005–18 average).



(e) Cardinal regions of Greater Santiago (total answers by quarter, 2005–18 average).



(f) Economic sectors (“Working” answers, 2005–18 average).

Source: Author’s calculations based on *Centro de Microdatos* database.

**Figure 1.** Actual inflation time series and considered attributes of the IEE.

**Table 1**

Re-coding key of considered attributes.

**1. Gender**

Original: Women, Men.

Re-coding: No re-coding.

**2. Age**

Original: 1=[0,14], 2=[15,24], 3=[25,34], 4=[35,44], 5=[45,54], 6=[55,65], 7=[65+].

Re-coding: 1=[0,34], 2=[35,65], 3=[65+].

**3. Education**

Original: 0=No education, 1=Elementary (incomplete), 2=Elementary (completed), 3=Secondary, 4=Technical, 5=College, 6=Other, 7=Non specified.

Re-coding: 1=[No education, Elementary, Secondary], 2=[Technical], 3=[College].

**4. County of living**

Original: 1=[Ñuñoa, La Reina, Macul, Peñalolén], 2=[San Miguel, La Cisterna, San Joaquín, La Granja, San Ramón, Pedro Aguirre Cerda, Lo Espejo], 3=[El Bosque, La Pintana, San Bernardo], 4=[Maipú, Cerrillos, Pudahuel, Lo Prado, Cerro Navia], 5=[Recoleta, Independencia, Conchalí, Renca, Quilicura, Huechuraba], 6=[Providencia, Vitacura, Las Condes, Lo Barnechea], 7=[Santiago, Estación Central, Quinta Normal], 8=[La Florida, Puente Alto].

Re-coding: 1=[San Miguel, La Cisterna, San Joaquín, La Granja, San Ramón, Pedro Aguirre Cerda, Lo Espejo, El Bosque, La Pintana, San Bernardo, Santiago, Estación Central, Quinta Normal], 2=[Ñuñoa, La Reina, Macul, Peñalolén, La Florida, Puente Alto], 3=[Providencia, Vitacura, Las Condes, Lo Barnechea], 4=[Maipú, Cerrillos, Pudahuel, Lo Prado, Cerro Navia, Recoleta, Independencia, Conchalí, Renca, Quilicura, Huechuraba].

**5. Economic sector [label]**

Original: 1=Agriculture, hunting, forestry, and fishing [Agr.,hun.,for.,&amp;fish.], 2=Mining and quarrying [Min.&amp;qua.], 3=Manufacturing industry [Manufacturing], 4=Construction [Construction], 5=Retail [Retail], 6=Government and financial services [Gov.&amp;fin.sv.], 7=Personal and household services [Per.&amp;hh.sv.], 8=Community and social services [Comm.&amp;soc.sv.], 9=Transportation, storage, communication, and public utility services [Tr.,st.,comm.,&amp;pub.ut.sv.].

Re-coding: No re-coding.

*Source:* Author's calculations.

a more meaningful result from where it is possible to draw some economics-based conclusions.<sup>9</sup> However, given the extensive answer classification in some of these questions, some factors deliver empty entries in quarters where no individuals fulfil the classification. To avoid this problem, I re-code some variables according to the key of Table 1.

The detailed definition of each economic sector is found in *Centro de Microdatos* (2009). The main differences between the two more apparently similar economic sectors are worth mentioning: *Personal and Households Services* and *Community and Social Services*. While repair services compound the former, laundries and laundry services, cleaning and dyeing establishments, domestic services, and personal and miscellaneous services (e.g., hairdressers and beauty salons, photo studios), the latter includes sanitation, educational, and health services, social welfare institutions, entertainment and leisure services, and other communal and social services (e.g., retail, professional, and labour associations; religious organisations).

As mentioned above, I consider all “*Working*” answers to the “*Employment situation*” question,

<sup>9</sup>This could be possible if, for instance, the economic sector in which consumers are currently working is more or less exposed to the business cycle. Thus an “exposure” mechanism to more relevant economic information could lead to more accurate forecasts (Faure and Medel, 2020).



corresponding to 59.1% of total answers on average, 2005–18 (see [Figure 1b](#)). The second and third place in terms of several answers are “*Housework*” and “*Others*”, from which are more difficult to delve further into their characteristics leading to more accurate forecasts as, for instance, considering the economic sector in which they worked (if applicable). However, more prohibitive is the sample availability to the inflation expectation question, impeding arriving at a reliable factor to be compared to those under the “*Working*” category.

In terms of gender and age, [Figure 1c](#), suggests that the (34,55) years old (re-coded) tranche is the biggest for both men and women, whereas the survey is biased towards women in all the original tranches. This fact could be potentially problematic, but all other attributes could partially correct this bias by refining the groups with a sample closer to the population.

Regarding education, [Figure 1d](#), display the distribution of the three tranches of education, re-coded to strengthen inflation factors with a more aggregate definition of “education”. Notice that they are displayed using the original eight zones in which the IEE is compiled, showing a greater concentration of a higher college degree level in the North-eastern part of Santiago and lower levels of education living in the Western part of the city. Aiming to strengthen factors with numerous observations, the eight zones of Santiago defined and surveyed by the IEE are re-coded following a cardinal representation: South-western, South-eastern, North-eastern, and North-western. According to [Figure 1e](#), cardinal zones are balanced in terms of total answers, except for the North-eastern zone. However, the correct representation is ensured through the different attributes shaping inflation factors, and then all of them are evaluated in the same manner. Finally, the sample distribution across the economic sectors is presented in [Figure 1f](#). There is no re-coding for this variable and thus, *Retail* (2005–18 average: 21.1%), *Community and Social Services* (18.9%), *Government and Financial Services* (18.5%), *Manufacturing* (13.8%), and *Personal and Household Services* (10.1%) are the biggest sectors with double-digit weight in the total sample of the IEE.

Despite the re-coding of some variables to strengthen the factors, it is still possible to have some of them with no entry for some quarters. Therefore, I use a criterion of each of the 648 factors disposing of at least 95% of the total possible observations (50 out of 55 observations). This criterion leaves out 55 of the 648 total factors that are listed in [Figure B.1](#) in [Appendix B](#). Also, given the reduced number of factors, it is possible to proceed with a brute-force exercise testing all factors instead of an algorithm-based search for the attributes leading to the best out-of-sample results.

## 2.2 Forecast Evaluation Framework

The forecast evaluation statistic used is the root mean squared forecast error (RMSFE), defined as:

$$RMSFE = \left[ \frac{1}{P} \sum_{t=1}^{T=55} (\pi_{\tau+4} - \pi_{\tau+4|t})^2 \right]^{\frac{1}{2}}, \quad (1)$$

where  $\pi_{\tau+4|t}$  represents the transformed 4-quarter-ahead forecast of  $\pi_{\tau+4}$  made with information known until time  $t$ . I dispose of a total of  $P = 55$  forecasts, ranging from 2005.II to 2018.IV in quarterly frequency. For simplicity, the results are reported using a relative measure of the

RMSFE, easing a comparison across the alternative forecasts:

$$\text{RMSFE ratio} = \text{RMSFE}^{\text{Factor}} / \text{RMSFE}^{\text{RW}}, \tag{2}$$

where “Factor” stands for each of the 55 available factors; thus, values below one imply a better performance in favour of the consumer-based factor.

To investigate to what extent the predictive gains are statistically significant, I make use of the unconditional one-sided *t*-type [Giacomini and White \(2006; GW\)](#) test providing the advantage of comparing forecasting methods instead of forecasting models. As the null hypothesis (NH) is defined as *the competing forecast has a superior predictive ability compared to the RW*, a one-sided *t*-type GW statistic is used accordingly.

Formally, I test the NH:  $\mathbb{E}_\tau(d_\tau) \leq 0$  against the alternative AH:  $\mathbb{E}_\tau(d_\tau) > 0$ , where

$$\widehat{d}_\tau = \left(\pi_{\tau+4} - \pi_{\tau+4|\tau}^{\text{RW}}\right)^2 - \left(\pi_{\tau+4} - \pi_{\tau+4|\tau}^{\text{Factor}}\right)^2, \tag{3}$$

using the [Newey and West \(1987\)](#) heteroskedasticity and autocorrelation corrected estimator of the standard deviation of  $\widehat{d}_\tau$ . 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  $\alpha\%$ .

### 3. Results

#### 3.1 RMSFE Results

The RMSFE ratio results are presented in [Table 2](#) for both the end-of-period and the average high-to-low frequency transformation. As mentioned above, the figures below one favour the consumer-based inflation factor. Factors are coded with the mask “[Gender]-[Age group]-[Education]-[Cardinal region]-[Economic sector]” and shown in lexicographic order. Also, those coming out better than the RW are labelled as *Fj*, where *j* is the ranking according to the RMSFE results as shown in [Table 3](#). A further visual comparison can be found in [Appendix C](#).

**Table 2**  
RMSFE ratio and Giacomini-White test results.

Factor	End-of-period	Average
M-[.,34]-Secondary-NWest-Retail	3.307**	3.404**
M-[.,34]-College-SEast-Gov. & fin. sv.	1.880	1.967
M-[.,34]-College-NEast-Gov. & fin. sv.	1.054	1.058
M-[35,65]-Secondary-SWest-Manufacturing	1.654	1.691
M-[35,65]-Secondary-SWest-Construction	1.303	1.333
M-[35,65]-Secondary-SWest-Retail	2.062**	2.115**
M-[35,65]-Secondary-SWest-Gov. & fin. sv.	2.424**	2.490**
M-[35,65]-Secondary-SWest-Tr., st., comm., & pub. ut. sv.	2.387	2.479
M-[35,65]-Secondary-SEast-Manufacturing	1.105	1.131
M-[35,65]-Secondary-SEast-Construction	1.587*	1.641*
M-[35,65]-Secondary-SEast-Retail	1.383	1.387
M-[35,65]-Secondary-SEast-Gov. & fin. sv.	1.262	1.311
M-[35,65]-Secondary-SEast-Per. & hh. sv.	2.717	2.752
M-[35,65]-Secondary-SEast-Tr., st., comm., & pub. ut. sv.	1.927*	1.964*

(Continued on next page)

Table 2 (continued)

Factor	End-of-period	Average
M-[35,65]-Secondary-NWest-Manufacturing	1.335	1.365
M-[35,65]-Secondary-NWest-Construction	1.998	2.008
M-[35,65]-Secondary-NWest-Retail	1.482	1.509
M-[35,65]-Secondary-NWest-Gov. & fin. sv.	2.342	2.375
M-[35,65]-Secondary-NWest-Retail	1.268	1.289
M-[35,65]-Secondary-NWest-Tr., st., comm., & pub. ut. sv.	1.410	1.421
M-[35,65]-College-SWest-Retail	1.121	1.162
M-[35,65]-College-SWest-Gov. and fin. sv.	0.912	0.914
M-[35,65]-College-SWest-Retail	0.956	0.966
M-[35,65]-College-SEast-Retail	1.829	1.885
M-[35,65]-College-SEast-Gov. & fin. sv.	0.913	0.910
M-[35,65]-College-SEast-Retail	0.862	0.865
M-[35,65]-College-NEast-Retail	2.071	2.096
M-[35,65]-College-NEast-Gov. & fin. sv.	0.862	0.856
M-[35,65]-College-NEast-Retail	0.902	0.895
M-[35,65]-College-NWest-Gov. & fin. sv.	0.898	0.885
W-[.,34]-Secondary-SEast-Retail	2.189**	2.267***
W-[.,34]-Secondary-NWest-Retail	1.479**	1.517*
W-[.,34]-College-SEast-Retail.	2.137	2.183
W-[35,65]-Secondary-SWest-Manufacturing	2.253***	2.327***
W-[35,65]-Secondary-SWest-Retail	1.974*	2.015*
W-[35,65]-Secondary-SWest-Per. & hh. sv.	2.586***	2.666***
W-[35,65]-Secondary-SWest-Retail	1.642*	1.674*
W-[35,65]-Secondary-SEast-Manufacturing	2.556	2.608
W-[35,65]-Secondary-SEast-Retail	2.257**	2.303**
W-[35,65]-Secondary-SEast-Gov. & fin. sv.	1.976	2.016
W-[35,65]-Secondary-SEast-Per. & hh. sv.	2.246*	2.264*
W-[35,65]-Secondary-SEast-Retail	2.369	2.409
W-[35,65]-Secondary-NWest-Manufacturing	2.110**	2.143**
W-[35,65]-Secondary-NWest-Retail	2.027*	2.081*
W-[35,65]-Secondary-NWest-Gov. & fin. sv.	1.525	1.546
W-[35,65]-Secondary-NWest-Per. & hh. sv.	1.437	1.453
W-[35,65]-Secondary-NWest-Retail	2.573	2.651
W-[35,65]-College-SWest-Retail	1.535	1.553
W-[35,65]-College-SEast-Retail	1.769	1.801
W-[35,65]-College-NEast-Gov. & fin. sv.	1.031	1.018
W-[35,65]-College-NEast-Retail	0.811	0.826

Notes: Factor key: [Gender]-[Age group]-[Education]-[Cardinal region]-[Economic sector]; see Table 1. Shaded cells = RMSFE ratio < 1. Giacomini and White (2006) test's  $p$ -value: (\*\*\*)  $p < 1\%$ , (\*\*)  $p < 5\%$ , (\*)  $p < 10\%$ . Source: Author's calculations.

A remarkable result is that eight factors outperform the RW with both end-of-period and average versions, even in a similar (but not equal) ranking. In both cases, the winning factor comprises women aged between 35 and 65 years old, with a college degree, living in the North-eastern part of the city, and working in the *Community and Social Services* sector. This group is composed of up to 26 consumers, that is, the top 1% of forecasters. For an overview of its performance, Figure 2 displays the scatter plot of the actual inflation series and the  $F1$  winner factor. It is important to note that all points far from the  $45^\circ$  line are those from the half of

**Table 3**

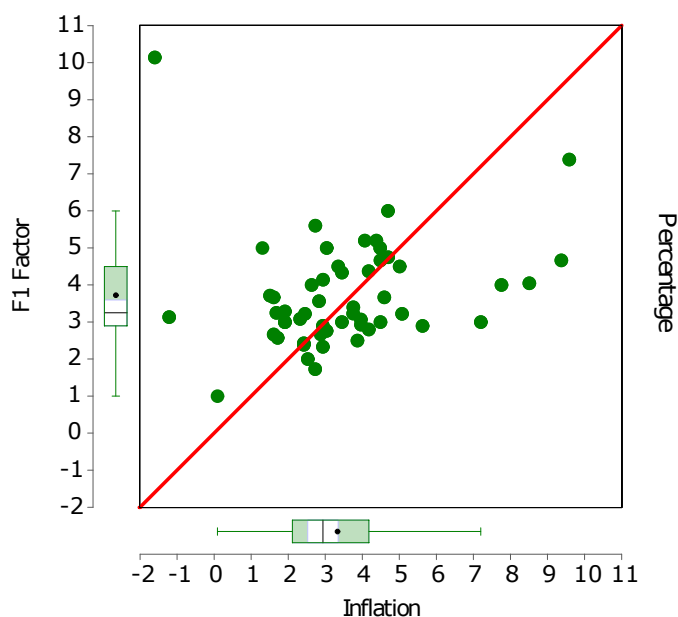
Inflation factors with a RMSFE smaller than the RW.

$F_i$	Factor	End-of-period	Average
$F_1$	W-[35,65]-College-NEast-Retail	0.811	0.826
$F_2$	M-[35,65]-College-SEast-Retail	0.862	0.865
$F_3$	M-[35,65]-College-NEast-Gov. & fin. sv.	0.862	0.856
$F_4$	M-[35,65]-College-NWest-Gov. & fin. sv.	0.898	0.885
$F_5$	M-[35,65]-College-NEast-Retail	0.902	0.895
$F_6$	M-[35,65]-College-SWest-Gov. and fin. sv.	0.912	0.914
$F_7$	M-[35,65]-College-SEast-Gov. & fin. sv.	0.913	0.910
$F_8$	M-[35,65]-College-SWest-Retail	0.956	0.966

Notes: See notes to Table 2. Source: Author’s calculations.

2007 to the half of 2009 when inflation suffers the impact of greater volatility due to the 2008–09 *Global Financial Crisis*. Despite this episode, the factor tends to show a concentration of accuracy, particularly when inflation is close to the target. In the second place, and still for the end-of-sample transformation, comes the factor composed by men aged between 35 and 65 years old, with a college degree, in a tie living in the South-eastern and North-eastern part of the city but working in *Retail* and *Government and Financial Services* sectors, respectively.

Then, in the fourth place are men aged between 35 and 65 years old, with a college degree, living in the North-western part of the city, working on *Government and Financial Services*. From the fifth to eighth places, I found men aged between 35 and 65 years old as a common characteristic and concentrated in *Community and Social Services* and *Government and Financial Services*,



Sample: 2005.II-2018.IV (55 observations). Source: Author’s calculations.

**Figure 2.** Actual inflation (e-o-p) and the best inflation factor scatter plot.

but with an uneven spatial distribution. No other than that, the mentioned eight factors come out as superior to the RW. The same ranking is observed for the average transformation except for a swap between the sixth and the seventh place.

According to the Giacomini-White test, a major drawback of these results is that none of the mentioned eight factors prove statistically superior to the RW. Notice that among the common characteristics of the eight best factors is that all of them are composed by consumers aged [35,65] and with a college degree. They are also composed of men, except for the winner factor. Regarding the economic sectors, despite being the option with more alternatives available, the best results are concentrated in just two sectors: *Government and Financial Sector* and *Community and Social Services*. Finally, there is no clear pattern regarding the spatial location.

### 3.2 Complementary Econometric Results

After finding the most promising factors, that is, those exhibiting a better forecasting performance than the RW (Table 3), I conduct econometric exercises to further discriminate between them. The first exercise is the simplest one and compares practical-use descriptive statistics. The results are presented in Table 4. Factors closest to the total end-of-period mean are *F3*, *F5*, and *F6*, whereas for the median is *F5*, *F3*, and *F1*. When comparing the standard deviations, *F3* and *F5* again come out with satisfactory results, replicated to the percentage of times within the (2,4) interval. In this sense, *F3* and *F5* display similar in-sample diagnostic as total end-of-period inflation, but it is relevant to delve into their predictive features with the remaining exercises.

**Table 4**

Complementary econometric results: descriptive statistics.

	Total	Total								
	E-o-p	Average	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	<i>F5</i>	<i>F6</i>	<i>F7</i>	<i>F8</i>
RMSFE ratio	-	-	0.811	0.862	0.862	0.898	0.902	0.912	0.913	0.956
Mean	3.335	3.428	3.754	4.025	3.490	3.798	3.322	3.729	3.962	3.953
Median	2.932	3.176	3.268	3.481	3.268	3.667	3.091	3.310	3.586	3.500
Std. deviation	2.122	2.150	1.467	1.808	1.282	1.771	1.235	2.297	2.263	2.430
% within (2,4)	65.9%	49.1%	59.2%	62.3%	66.0%	43.1%	74.5%	46.7%	58.5%	42.2%

Notes: *F1*...-*F8* are the eight factors highlighted in Table 2 sorted according to its RMSFE ratio as in Table 3.

Sample: 2005.II-2018.IV (55 observations). Source: Author's calculations.

The second exercise compares the biases of the forecasting series. Expectations are unbiased (i.e., they show no room for improvement) if they are on average equal to actual inflation. This translates into the regression:

$$\pi_{\tau+4|\tau-4} = \alpha + \beta\pi_{\tau|\tau} + \varepsilon_{\tau}, \quad (4)$$

where  $\pi_{\tau+4|\tau-4}$  is the 12-month-ahead forecast (but transformed to quarters), lagged in four quarters to be comparable to actual inflation  $\pi_{\tau|\tau}$ ,  $\alpha$  and  $\beta$  are parameters to be estimated, and  $\varepsilon_{\tau}$  is an error term that could be autocorrelated, in whose case a 3-term moving average is used.

**Table 5**

Complementary econometric results: unbiasedness.

	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	<i>F5</i>	<i>F6</i>	<i>F7</i>	<i>F8</i>
Equation: $\pi_{\tau+4 \tau} = \alpha + \beta\pi_{\tau} + \varepsilon_{\tau}$								
$\alpha$	<b>3.515</b>	<b>3.222</b>	<b>2.600</b>	<b>2.951</b>	<b>2.811</b>	<b>2.526</b>	<b>2.469</b>	<b>2.996</b>
<i>p</i> -value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.007)	(0.007)
$\beta$	0.067	0.232	<b>0.258</b>	<b>0.243</b>	<b>0.149</b>	<b>0.339</b>	0.432	0.280
<i>p</i> -value	(0.653)	(0.150)	(0.000)	(0.070)	(0.036)	(0.000)	(0.113)	(0.371)
<i>F</i> -test ( $\alpha = 0, \beta = 1$ )	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.017
<i>p</i> -value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.011)
Adj. R-sq.	0.000	0.062	0.181	0.075	0.056	0.079	0.161	0.033

*Notes:* *F1*–...–*F8* are the eight factors highlighted in Table 2 sorted according to its RMSFE ratio as in Table 3. Bold estimates show statistically significant coefficients at 10% level of confidence. *p*-value shown in parentheses. “Adj. R-sq.” stands for adjusted goodness-of-fit coefficient. “*F*-test ( $\alpha = 0, \beta = 1$ )” stands for the *F*-statistic of the null hypothesis  $\alpha = 0$  and  $\beta = 1$ .

*Sample:* 2005.II-2018.IV (55 observations). *Source:* Author’s calculations.

The frequency transformation implies replacing the  $t + 12$  subscript for  $\tau + 4$ . So, the lack of bias simultaneously means that  $\alpha = 0$  and  $\beta = 1$ , which is tested through a Wald-type *F*-test. If the null hypothesis NH:  $\alpha = 0, \beta = 1$  is rejected, then the forecast is biased, and there is room for improvements; otherwise, the prediction is said to be *efficient* in using the information available up to the release of the latest observation.

The results are presented in Table 5. The results suggest that the null hypothesis of unbiasedness is rejected for all factors and thus, there is room for improvement in all of them. However, the results of factors *F1* and *F2* suggest that they are more efficient in using information because only the statistically significant coefficient is unrelated to actual inflation. Thus, for instance, a less-invasive intercept-correction method could deliver better results than the remaining factors, which requires more information to achieve efficiency. This is the case of the *F3* and *F5* factors, in which  $\beta$  result as statistically significant—0.258 and 0.149, respectively—and thus, they are not considering the information of actual inflation when estimated. Notice, however, that this evaluation of factors is made to identify the characteristics of the group that better forecast inflation and not entirely of forecast accuracy, in the sense that exploring more with these factors—particularly with these with a good behaviour—could deliver the best results after several statistical treatments and factor combinations, a task left for further research.

The third exercise, which encompasses the results of bias when present, consists in asking the degree to which forecast variables are forward-looking versus backward-looking. A static version of this test is represented with the regression:

$$\pi_{\tau+4|\tau} = \lambda + \gamma\pi_{\tau+4} + (1 - \gamma)\pi_{\tau-1} + v_{\tau}, \tag{5}$$

where  $\lambda$  and  $\gamma$  are parameters to be estimated, and  $v_{\tau}$  is an error term that could be autocorrelated, in whose case a 3-term moving average is used. Thus, as the  $\gamma$ -parameters are constrained to add to unity, a relatively greater  $\gamma$  parameter reflects a higher degree of forward-lookingness

**Table 6**

Complementary econometric results: forward lookingness.

	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	<i>F5</i>	<i>F6</i>	<i>F7</i>	<i>F8</i>
Equation: $\pi_{\tau+4 \tau} = \lambda + \gamma\pi_{\tau+4} + (1 - \gamma)\pi_{\tau-1} + v_{\tau}$								
$\gamma$	<b>0.436</b>	<b>0.330</b>	<b>0.420</b>	<b>0.315</b>	<b>0.454</b>	<b>0.343</b>	0.215	0.218
<i>p</i> -value	(0.001)	(0.003)	(0.000)	(0.008)	(0.000)	(0.001)	(0.137)	(0.241)
$1 - \gamma$	<b>0.564</b>	<b>0.670</b>	<b>0.580</b>	<b>0.685</b>	<b>0.546</b>	<b>0.657</b>	0.785	0.782
<i>p</i> -value	(0.001)	(0.003)	(0.000)	(0.008)	(0.000)	(0.001)	(0.137)	(0.241)
Adj. R-sq.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

*Notes:* *F1*-...-*F8* are the eight factors highlighted in Table 2 sorted according to its RMSFE ratio as in Table 3. Bold estimates show statistically significant coefficients at 10% level of confidence.

*p*-value shown in parentheses. “Adj. R-sq.” stands for adjusted goodness-of-fit coefficient.

*Sample:* 2005.II-2018.IV (55 observations). *Source:* Author’s calculations.

of the factor, which is desirable for a forecasting variable. As Łyziak (2013) posits, a  $\gamma = 1$  parameter suggests that inflation forecasts are fully forward-looking and meet the requirement that the rational expectations hypothesis be unbiased. In contrast, if  $\gamma$  is not statistically different from zero, inflation forecasts are fully backward-looking, being very easy to outperform and, thus, providing very little informational content. The results are presented in Table 6. In this sense, *F1* and *F5* factors come out as the best options showing a share close to 50% of forward-lookingness. Remarkably, neither of the estimated coefficients comes out significant for *F7* and *F8* factors. This result is unexpected given their satisfactory results in the rest of the exercises. It is argued that, so far, *F1* (and recalling that it is the best factor according to the RMSFE ratio) and *F5* are the factors with the desirable features of a forecasting variable, and that the only difference in their composition is that while *F1* is compounded by women, *F5* is compounded by men.

Much empirical research has been conducted on the Phillips curve regarding the fourth exercise. Particularly after the proposal of Galí and Gertler (1999) of the hybrid New Keynesian Phillips curve including direct measures of inflation expectations, many authors have fit this type of Phillips curve for forecasting purposes.<sup>10</sup> The success of this version of the Phillips curve relies on the mixture of two features of price dynamics, namely, its persistence, captured with lagged inflation, and the prospective price formation by firms captured by direct measures of expectations. This is added to a cost-push measure such as the output gap, reflecting inflationary pressures from the real economy. Thus, if inflation expectations act as so, they must be statistically significant in this setup, as argued in Łyziak (2013). Thus, the fourth exercise consists in using each promising inflation factor in a hybrid New Keynesian Phillips curve described as:

$$\pi_{\tau} = \bar{\pi} + \rho\pi_{\tau-1} + (1 - \rho)\pi_{\tau+4|\tau} + \phi\tilde{y}_{\tau} + \psi_{\tau}, \quad (6)$$

<sup>10</sup>See, for instance, Paloviita and Mayes (2005) for 11 European countries, Nason and Smith (2008) for the United States, Henzel and Wollmershäuser (2008) for Italy, Paloviita (2009) for the Eurozone, Jean-Baptiste (2012) for the UK, and Medel (2015, 2017) and Marcel et al. (2017) for the case of Chile.

**Table 7**

Complementary econometric results: hybrid New Keynesian Phillips curve.

	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	<i>F5</i>	<i>F6</i>	<i>F7</i>	<i>F8</i>
Phillips curve: $\pi_\tau = \bar{\pi} + \rho\pi_{\tau-1} + (1 - \rho)\pi_{\tau+4 \tau} + \phi\tilde{y}_\tau + \psi_\tau$								
$\rho$	<b>0.884</b>	<b>0.836</b>	<b>0.772</b>	<b>0.866</b>	<b>0.794</b>	<b>0.810</b>	<b>0.790</b>	<b>0.894</b>
<i>p</i> -value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$1-\rho$	<b>0.116</b>	<b>0.164</b>	<b>0.228</b>	<b>0.134</b>	<b>0.206</b>	<b>0.190</b>	<b>0.210</b>	<b>0.106</b>
<i>p</i> -value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\phi$	<b>0.304</b>	<b>0.316</b>	<b>0.254</b>	<b>0.281</b>	<b>0.285</b>	<b>0.241</b>	<b>0.336</b>	<b>0.407</b>
<i>p</i> -value	(0.034)	(0.032)	(0.050)	(0.047)	(0.000)	(0.088)	(0.011)	(0.028)
Adj. R-sq.	0.705	0.721	0.740	0.718	0.737	0.732	0.745	0.683
RMSFE ( <i>h</i> = 4)	1.768	1.404	1.501	1.725	1.509	1.791	1.440	2.085

*Notes:* *F1*–...–*F8* are the eight factors highlighted in Table 2 sorted according to its RMSFE ratio as in Table 3. Bold estimates show statistically significant coefficients at 10% level of confidence. *p*-value shown in parentheses. “Adj. R-sq.” stands for adjusted goodness-of-fit coefficient. All estimates are performed with the Ordinary Least Squares method, and the Newey and West (1987) heteroskedasticity and autocorrelation corrected estimator for the standard deviation. For forecasting purposes, the sample 2005.III–2012.IV (30 observations) is used for estimation, whereas the sample 2013.I–2018.IV (24 observations) is used for evaluation. *Sample:* 2005.II–2018.IV (55 observations). *Source:* Author’s calculations.

where  $\tilde{y}_\tau$  is the output gap,<sup>11</sup>  $\bar{\pi}$ ,  $\rho$ , and  $\phi$  are parameters to be estimated, and  $\psi_\tau$  is an error term that could be autocorrelated, in whose case a 3-term moving average is used. The criteria to determine which inflation factor is preferable are based on both the statistical significance of the associated coefficient and the improvements in the model’s goodness-of-fit. A third criterion is based on the 4-quarter-ahead forecast obtained with each estimated version of equation (6) and evaluated according to a symmetric cost function (the RMSFE). The estimation sample considers the first 30 observations for estimation (2005.III–2012.IV, 30 observations), leaving the remaining part of the sample (2013.I–2018.IV, 24 observations) for forecast evaluation. The results are presented in Table 7. It is observed that all factors display statistically significant results and, thus, the discriminatory power seems low. However, the differences are more noticeable when making predictions with these estimates (actual end-of-period inflation four quarters ahead). In particular, *F2* and *F7* factors display the lowest RMSFE, followed by the *F3* and *F5* factors, and then *F4* and *F1* factors. Thus, the hybrid New Keynesian Phillips curves based on neither *F1* nor *F5* are on the top three of inflation forecasts, with the caveat that no statistical inference is carried out between them. These results could be explained by a slightly greater output gap coefficient allowed by the factors leading to the best forecasting results and the difficulty of

<sup>11</sup>The approximation used here for the output gap is obtained as the difference between the logarithmic level of actual GDP and potential GDP. The latter is defined as the logarithmic level of the GDP’s seasonally adjusted and filtered version including up to five years of forecast observations coming from an *ad-hoc* autoregressive integrated moving average (ARIMA) model. This last step is performed to avoid the “end-of-sample” identification problem when using the Hodrick-Prescott ( $\lambda = 1,600$ ) method to filter the series. The seasonal adjustment program used is the X-13ARIMA-SEATS, whereas the ARIMA forecasting model is the so-called *airline model* (Box and Jenkins, 1970; Ghysels et al., 2006).



estimating coefficients with covariates more correlated between them.

Finally, the fifth exercise consists of the so-called U-Theil decomposition of forecast errors in bias, regression, and disturbance proportions. This decomposition aims to disentangle and provide a taxonomy of forecast errors aiming to determine their sources. Consequently, depending on the source, an improvement strategy could be deployed. As [Ahlburg \(1984\)](#) states, following [Theil \(1971\)](#), the decomposition of the mean squared forecast error (MSFE) is:

$$MSFE = \frac{1}{P} \sum_{\tau=1}^P (\pi_{\tau+4|\tau} - \pi_{\tau+4})^2 = \underbrace{(\bar{\pi}_{\tau+4|\tau} - \bar{\pi}_{\tau+4})^2}_{\text{Bias proportion}} + \underbrace{(\sigma_{\pi_{\tau+4|\tau}} - \rho\sigma_{\pi_{\tau+4}})^2}_{\text{Regression proportion}} + \underbrace{(1 - \rho^2)\sigma_{\pi_{\tau+4}}^2}_{\text{Disturbance proportion}}, \quad (7)$$

where  $\bar{\pi}_{\tau+4|\tau}$  and  $\bar{\pi}_{\tau+4}$  are the means of the predicted and actual values of inflation series, respectively, whereas  $\sigma_{\pi_{\tau+4|\tau}}$  and  $\sigma_{\pi_{\tau+4}}$  are their standard deviations, and  $\rho$  is the correlation coefficient between the predicted and actual values of the series. The bias proportion arises from the systematic under- or over-estimation of the mean of the target variable. In contrast, the regression proportion is due to the slope coefficient obtained of the relationship between the actual value of the series and its forecast value. These two sources of forecasting errors are systematic and could be reduced to improve accuracy. In contrast, the disturbance proportion is the non-systematic error, not allowing a role for the forecaster. To ease comparison between these three components, they are re-scaled to the MSFE to be now comparable across inflation factors. A factor with the highest disturbance proportions is, thus, a signal that forecasters are more efficient in the use of information. Notice that this evaluation is made with inflation factors obtained as they are, without any correction or combination. This is relevant as some factors that exhibit a worse performance than the RW in the first step, in an eventual second step of correction, may outperform some factors with promising initial accuracy. This kind of analysis—a data mining with forecasting purposes—while useful certainly goes beyond the scope of this article. Also, a combination or accuracy enhancing technique will blur the true contribution of a specific group to forecast accuracy, being difficult to identify.

The results are presented in [Table 8](#). As mentioned above, a greater share of the disturbance proportion is desirable from the standpoint of this evaluation, as it aims to reveal how efficient the respondents are in the use of information and expectation formation without any *ex-post* statistical intervention. Consequently, the *F1* factor comes out as the best alternative because it shows the largest share of disturbance and, by construction, the smallest share of bias and noise due to the regression. In these terms, the *F5* factor has the best options to be corrected as displaying the highest regression proportion.

A final in-depth analysis of *F1* and *F5* factors consists in hand-picking observations in which they both perform poorly and transform them into missing observation. On the one hand, for the *F1* factor, an observation with a greater deviation is found in 2009.IV (showing 6.0% when the actual inflation was -2.1%). Dropping this observation reduces the RMSFE ratio from 0.811 to 0.807. On the other hand, for the *F5* factor, two major deviations are noticed; one in 2009.IV (6.0% versus -2.1% actual) and another in 2010.IV (7.5% versus 2.0% actual). Dropping these observations implies dropping the RMSFE ratio from 0.902 to 0.878—thus, a fall in the RMSFE

**Table 8**

Complementary econometric results: U-Theil.

	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	<i>F5</i>	<i>F6</i>	<i>F7</i>	<i>F8</i>
U-Theil decomposition: Squared errors = Bias + Regression + Disturbance								
Bias	2.4%	8.1%	0.1%	2.0%	0.1%	2.0%	22.3%	13.5%
Regression	17.3%	17.4%	67.7%	50.6%	87.4%	43.9%	10.5%	21.8%
Disturbance	80.3%	74.6%	32.1%	47.3%	12.5%	54.1%	67.2%	64.6%

*Notes:* *F1*-...-*F8* are the eight factors highlighted in Table 2 sorted according to its RMSFE ratio as in Table 3. Each column adds to unity.

*Sample:* 2005.II-2018.IV (55 observations). *Source:* Author’s calculations.

ratio is not enough to outperform the *F1* factor. Therefore, the *F1* factor still stands as the best factor. Also, a remarkable fact is that during the 2008–09 *Global Financial Crisis*, the *F1* factor recorded an inflation forecast of 10.1% while the effective rate was 9.9%. At the same time, the *F5* factor registered a wider difference by recording a rate of 4.8%.<sup>12</sup>

In sum, according to the analysis of complementary econometric exercises, the *F1* factor (*W*-[35,65]-*College-NEast-Retail*) is consistently the best in terms of accuracy and desired features expected from a forecasting variable. However, the *F5* factor (*M*-[35,65]-*College-NEast-Retail*), despite not being the second-best in terms of accuracy measured through the RMSFE ratio, comes out as a valid option fulfilling the behaviour of a forecasting variable.

#### 4. Discussion and Directions for Further Research

In this section, I analyse three issues of interest in the light of results: (i) the extent to which they have common features with the international evidence, (ii) an explanation on why I use the mean instead of the median as the statistic that built inflation factors, and (iii) an exploratory analysis on the results of the less accurate factor when forecasting total inflation (*M*-[.,34]-*Secondary-NWest-Retail*, see Table 2), questioning which granular prices respondents may be targeting at when they are surveyed.

The first issue of this section is to analyse to what extent the characteristics of the best forecasters found in this paper relates to the international evidence. Three facts are worth mentioning: (i) the results of this paper could be easily summarised as higher-income (due to its county of living) women aged between 35–65 years old with a college degree are the best forecasters within the survey (when using the whole sample), (ii) it is very common to find that higher-income households, men, more educated, and older respondents are the best forecasters when analysing international surveys, and (iii) it is very uncommon to find women as the best forecasters.

<sup>12</sup>A final exploratory check is regarding forecast combinations. An exercise combining both *F1* and *F5* factors with linear weights adding to unity suggest that pairwise combinations do not improve the accuracy of the *F1* factor. In particular, using the 9-ordered-pair grid ranging from 0.1 to 0.9 as weight for *F1* (and, consequently, of 0.9 to 0.1 for *F5*) deliver RMSFE ratios of 0.890, 0.874, 0.860, 0.849, 0.839, 0.831, 0.826, 0.823, and 0.823 which are all below the 0.811 RMSFE ratio exhibited by the *F1* factor alone.

The evidence collected from the *Survey of Consumers* of the University of Michigan for the United States gives a good benchmark. It reports the personal characteristics of the respondents, such as gender, age, marital status, education, race, and income level. The findings generally suggest that white men with more than high-school educational level, within the 55–65 years-old range provide more accurate inflation forecast 12-month-ahead (Bryan and Venkatu, 2001a,b; Souleles, 2004; Pfajfar and Santoro, 2008; Meyer and Venkatu, 2011; Madeira and Zafar, 2015). The “race” variable is included in Bryan and Venkatu (2001a), Meyer and Venkatu (2011), Madeira and Zafar (2015), Axelrod et al. (2018) when using the *Survey of Consumers*, and Rossouw et al. (2011) analysing the South African case, playing a role in forecast accuracy. However, it is not captured by the IEE survey as Chile’s Metropolitan Area had a relatively low race diversity when the survey’s sampling was set. Nowadays, however, as the Central Bank of Chile (2018) points out, an immigration shock was experienced during 2015–17 in Chile which was of a substantial number of people (achieving 8.8% of the labour force according to the 2017 Census) and concentrated in the Metropolitan Area. Also, the origin countries and race of this immigration wave are more diverse and show a significant participation in the labour market (a rate of 77% registered in the March–May 2017 moving quarter). Thus, future sample updating of the IEE could include the race variable.

Also, Palmqvist and Strömberg (2004) using a survey conducted by *Statistiska Centralbyrån* (Statistics Sweden) and *Growth from Knowledge* (GfK) for Sweden, Armantier et al. (2017), using the New York Fed’s *Survey of Consumer Expectation* for the United States, Campelo et al. (2019) using the Brazilian IBRE/FGV *Consumer Confidence Survey*, among others cited above using the *Survey of Consumers* of the University of Michigan, includes geographic zones which are determinant of inflation expectations differences among demographic groups but, by nature, they are uncomparable across surveys.

Interestingly, Sabrowski (2008) uses the *Business and Consumer Survey* conducted by the European Commission for Germany to find a statistically significant role of labour status (working/unemployed) when finding the best forecasting demographic group. This issue was also treated in Malgarini (2009) for the Italian case. Similarly, Ehrmann et al. (2017) analyse the role of respondents’ financial situation using the *Survey of Consumers*, finding that more credit-constrained respondents tend to overestimate inflation rates. Goyal and Parab (2019a,b) include the assessment of the economic outlook of respondents when analysing consumers’ inflation expectations for the case of India. In contrast, Ichiue and Nishiguchi (2015) analyse the role of asset holdings and financial literacy in the case of Japan. Diamond et al. (2020) add the type of contract of the respondent as determinant of inflation expectations, also for Japan.

The overwhelming majority of survey-based results on inflation expectations by demographic groups suggest men as the best forecasters when controlling for all mentioned variables. This is the main difference of the results of this paper with the international evidence. In general, women tend to overestimate inflation because they put an additional attention to prices with a higher volatility, overreacting when forming their inflation expectations. This hypothesis was firstly analysed in Jonung (1981) for the case of Sweden and proposed in Pfajfar and Santoro (2008) using the *Survey of Consumers*. However, in Chile’s case, this fact seems to favour women when

forming their inflation expectations. Three reasons could yield this advantage: women spend more time screening a major portion of CPI-basket item prices, or the Chilean total inflation is specifically driven by those items in which women pay more attention than men (i.e., food stuff), as suggested by [Jonung \(1981\)](#), or a combination of both reasons.

Regarding the first reason, there is some supporting evidence in the Chilean *National Survey on Time Use* (ENUT) conducted by the [National Statistics Institute \(2015\)](#), indicating that within the group of employed, men dedicate an average of 2.85 hours a day to “unpaid work” (a category that includes purchases of home stuff), while women allocate an average of 5.85 hours a day to these tasks (difference women minus men: +3 hours). This difference grows within the unemployed, in which men spend an average of 3.49 hours a day, while women 7.11 hours (difference: +3.62 hours), and a slightly larger difference (+3.69 hours) in inactive people, with men allocating 2.54 hours, while women 6.23 hours, which is almost equal to an office-based working time of 8 hours per working day.

This explanation is also in line with [Reis’s \(2006a; 2006b\)](#) claim, suggesting that different agents of the economy have different incentives to “rational in-attend” changes in prices. This could explain the difference across income levels and, added to the results regarding time use; it also could explain the difference between men and woman. It is important to remark that, for the case of this study, the sample available to construct women and “*Housework*”-employment situation-based inflation factor does not allow to perform a reliable econometric analysis. Thus, it is impossible within this framework to test “*Working*”-woman against “*Housework*”-woman forecasting accuracy without using a more complex sampling method, a task left for further research.

The second issue of this section relates to the use of the mean instead of the median of each group of respondents as the statistic used to build factors. [Meyer and Venkatu \(2011\)](#) suggest that the appropriate statistic to represent and compare demographic groups in inflation expectations, should be the median instead of the mean. However, a first important difference of this study with [Meyer and Venkatu \(2011\)](#) is the sampling method, sample span and representativeness, and the frequency in which the survey is conducted. As [Meyer and Venkatu \(2011\)](#) use the *Survey of Consumers* elaborated by the University of Michigan, they dispose of a minimum of 500 answers per month with samples designed to represent all American households. This excludes those in Alaska and Hawaii, which differs to the Chilean case analysed in this paper, referring to the Greater Metropolitan Area (and the CPI still being elaborated countrywide). Also, the question about inflation expectations<sup>13</sup> has been asked since the early 1980s, thus, covering a period of 30 years at the time of that study (2011). Instead, the Chilean data of *Centro de Microdatos* covers a period of 13 years on a quarterly basis. All these differences pose a titanic challenge to ensure the right representation of each characteristic that together compound an inflation factor.<sup>14</sup>

Nevertheless, and even more relevant for the exercise carried out in this study, it is key that

<sup>13</sup>“During the next 12 months, do you think that prices, in general, will go up, or go down, or stay where they are now?”.

<sup>14</sup>[Kenny et al. \(2015\)](#) propose a complete setup making use of panel estimates using density forecasts of respondents of the European Central Bank’s *Survey of Professional Forecasters* at micro data level. This is an avenue to be explored for the case of Chile when a unique identificatory become available.

when more respondents fulfil the characteristics that define the factor, they can re-shape the distribution function and change the group's outcome: this occurs when using the mean instead of the median. This makes that any respondent that fulfils characteristics counts and, as lack of more information, any of them do that with the same weight  $1/N$  (assuming randomness with  $N$  being the number of respondents of the factor). In contrast, when using the median, respondents could be *irrelevant* despite fulfilling the criteria to be part of the factor, contradicting the setup of this exercise. For instance, a factor could be compounded of  $N - 2$  respondents, and two new respondents must be added to the factor. The factor displays a median of  $\bar{M}_{N-2}$  upon using the  $N - 2$  sample. Suppose the new entries are  $(\bar{M}_{N-2} + \varepsilon_1)$  and  $(\bar{M}_{N-2} - \varepsilon_2)$ , where  $0 < (\varepsilon_1, \varepsilon_2) < \gamma$  with  $\gamma \leq \min\{|x^{\max} - \bar{M}_{N-2}|, |\bar{M}_{N-2} - x^{\min}|\}$ , where  $x^{\max}$  and  $x^{\min}$  are the maximum and the minimum value of inflation expectation reported, then  $\bar{M}_{N-2} = \bar{M}_N$ , and the two new observations become irrelevant when considering the median—despite that the respondents fulfil the characteristics of the factor. Thus, it must be included. However, when using the mean, the case  $\bar{M}_{N-2} = \bar{M}_N$  only happens if  $\varepsilon_1 = \varepsilon_2$  and all respondents are included in any case. Also, if  $\varepsilon_1 > \varepsilon_2$  then  $\bar{M}_{N-2} < \bar{M}_N$ , and if  $\varepsilon_1 < \varepsilon_2$  then  $\bar{M}_{N-2} > \bar{M}_N$ , making every respondent to have an impact on the overall score of the factor. As the exercise carried out in this paper makes every respondent count (in a large extent because of the sampling method of the database), I opt to use the mean as the chosen point of the distribution.

In any case, the Table 1D in Appendix D display the results mimicking Table 2 but using the median instead of the mean. The results are sensitive to this change. A total of 14 (instead of 8) factors display an RMSFE ratio below unity, explained because of the insensitivity of the median to outliers. Yet, according to the Giacomini-White test, none of the factors comes out as statistically different from the RW. All of the eight factors that show a better performance than the RW using the mean are still better than the RW when using the median, except the *M-[35,65]-College-SWest-Retail* factor that is no longer better than the RW. This result highlights that excluding some respondents does not *always* lead to better results and choosing the point of the distribution function with the best forecasting results for a factor could be a more sophisticated task than calculating an automated statistic.

The third issue of this section consists of analysing the factor showing the most inaccurate performance when predicting total inflation (according to the RMFSE loss function) when forecasting the end-of-period headline inflation. The *M-[.,34]-Secondary-NWest-Retail* factor (see Table 2) shows an RMSFE ratio of 3.307. Thus, the question is whether the respondents with mentioned characteristics aimed at forecasting a specific price item or a reduced subset of the CPI basket. To that end, I use of the 2018 CPI basket data prices at the granular level constructed in Alvarado and Medel (2020) to calculate the correlation coefficient between mentioned factor and 303 items compounding the current CPI basket. An important feature is that the correlation is estimated contemporarily; without considering the time horizon in which the respondents give their answers and the horizon in which they are asked. This is so—and, thus, the used statistic is the correlation coefficient and not the RMSFE—because the respondents could easily think in “12-months ahead” as a vague future date without necessarily perfectly matching the months in which the questions are posed. This distinction is highly sensitive for the RMSFE calculation

**Table 9**

 CPI basket items with higher correlation ( $|40\%$ ) with the most inaccurate factor for total inflation.

No.	Item	Correlation of the item item with the most inaccurate factor	Correlation of the with the total CPI inflation	2018 CPI basket weight of the item
1	<i>Potatoes</i>	0.561	0.328	0.3362
2	<i>Canned vegetables</i>	0.505	0.294	0.0437
3	<i>Milk flavour</i>	0.441	0.282	0.0502
4	<i>Homeopathic medicines and food suppl.</i>	0.431	0.143	0.0855
5	<i>Furniture repair service</i>	0.400	0.147	0.0266

*Notes:* The most inaccurate factor when predicting total inflation is *M-[.,34]-Secondary-NWest-Retail*.

Items' time-series at granular level are obtained from [Alvarado and Medel \(2020\)](#).

*Sample:* 2010.I-2018.IV (36 observations). *Source:* Author's calculations.

over the 303 items of the CPI basket. Therefore, the results are exploratory in nature using the contemporary correlation coefficient.

The results are shown in [Table 9](#). These consist of all items with a correlation coefficient greater than  $|40\%$  with the most inaccurate factor (first column), obtaining five out of 303 items fulfilling this criterion;<sup>15</sup> almost all of them classified as necessities: potatoes, canned vegetables, milk flavour, homeopathic medicines and food supplements, and furniture repair service. Also, the correlation between the *M-[.,34]-Secondary-NWest-Retail* factor and total inflation is  $-0.11\%$  for the 2005.II–2018.IV (55 observations) sample, being almost unrelated between them in contemporaneous terms. To further delve into the target of the analysed factor, the items could be grouped in three sets: foodstuff, medicines, and home maintenance—which could be associated to a spending composition of lower-income households.

According to [CLAPES-UC \(2020\)](#), an estimation of the 2018-based CPI basket for the poorest-income quantile comes out with a greater weight for “*Food and non-alcoholic beverages*” of 26.34%, representing 7.04 basis points greater than the weight of the whole CPI basket. In contrast, “*Health*” represents 6.77% of the poorest-income quantile basket, -1 basis point less than that of the whole CPI basket, and “*Home equipment and maintenance*” comes out with 4.46% of the poorest-income quantile basket, corresponding to -2.07 basis points less than the weight in the whole basket. Note that a greater weight of food items is generally associated to poorer-income household spending. Also, as public health in Chile is free or provided with a subsidized low end-price for the poorer households, the CPI captures better private health prices that obey more to a market-oriented logic. Thus, it is not surprising that this category in the CPI less represent the poorest-income quantile. In addition, considering that the identified item corresponds to what is considered alternative medicine in Chile (homeopathic) rather than the conventional Western medicine, suggest cheaper treatment alternatives. Finally, it is indistinguishable from the label

<sup>15</sup>Notice that a higher threshold, e.g.,  $|50\%$ , leads to just two items (*Potatoes and Canned vegetables*) fulfilling the criterion. In contrast, lower thresholds, namely  $|30\%$ ,  $|20\%$ , and  $|10\%$  lead to 19, 45, and 78 items, respectively, being more difficult to find a common characteristic across them. So, a  $|40\%$  threshold led to a plausible number of items to be analysed.

“*Home equipment and maintenance*”, which share is associated to “*maintenance*”, but the spending in “*Furniture repair service*” certainly suggest the use of a second-best alternative to new home furniture. In sum, there is some evidence suggesting that the *M-[.,34]-Secondary-NWest-Retail* factor is inaccurate when forecasting total inflation because its respondents target a subset of the whole CPI. However, they could be very accurate when forecasting a CPI basket more oriented to lower-income households. Also, this result does not rule out that respondents compounding this factor follow a similar price-expectation formation process than those respondents of the most accurate factor but with different targets, which deserve more in-depth research.

## 5. Concluding Remarks

This article evaluated quantitative inflation forecasts for the Chilean economy, taking advantage of a specific survey of consumer perceptions at the individual microdata level, which, at the same time, is linked to a survey of employment and unemployment of Chile’s capital city, Santiago. Thus, the key advantage of the database is that it is possible to link, with no error, consumer perceptions and 12-month-ahead inflation forecasts with labour-market characteristics of the respondents, which heavily rely on personal (yet anonymous) characteristics. This is done by constructing sub-sets of inflation expectations factors with fully identifiable and mutually excluded characteristics such as gender, age, education, county of living, and the economic sector of present work.

By using a quarterly sample ranging from 2005.II to 2018.IV, the results for total inflation suggest that women aged between 35 and 65 years old, with a college degree, living in the North-eastern part of the city (that with the highest living standards of the country), and working in *Community and Social Services* sector are the best forecasters. Men aged between 35 and 65 years old, with a college degree, in a tie living in the South-eastern and North-eastern part of the city but working in *Retail* and *Government and Financial Services* sectors, respectively, are the second-best at forecasting inflation. Finally, men aged between 35 and 65 years old, with a college degree, living in the North-western part of the city, working on *Government and Financial Services* are in fourth place. From the fifth to eight places, I found men aged between 35 and 65 years old as a common characteristic and concentrated in *Community and Social Services* and *Government and Financial Services*, but with a different spatial distribution. These results are obtained by comparing mentioned inflation expectations factors to the naïve RW forecast. Only these eight out of 648 total possible factors outperform the RW, and none of them coming out as statistically superior according to the [Giacomini and White \(2006\)](#) test.

Several econometric exercises are also conducted to further discriminate between the best inflation factors, revealing that a factor different to the second-best in terms of forecast accuracy displays the characteristics required of a forecasting variable. Remarkably, this factor has the same specifications as the winner factor, with the only difference being that it is composed by men instead of women. Thus, potentially, there is space to delve into more intricate schemes to take full advantage of the predictive information of the overall survey—a task left for further research.

The results of this study are different to the general conclusions found in the international evidence, in the sense that women are better forecasters than men. All other comparable characteristics (higher income, higher educational level, and age range) align with the common findings. This led to think that, in Chile, either women pay more attention to a wider range of CPI-basket price items, or the Chilean total inflation is driven specifically by those items in which women form their expectations much better than men. Finally, the most inaccurate factor when predicting total inflation owes its result to the fact that it targets a subset of the CPI basket and focuses on a lower-income household spending. This does not necessarily suggest that they form their expectations in the wrong way or different from that of the winning factor.

It is important to remark that all this evaluation is made to reveal how efficient groups of respondents are in using information and inflation expectation formation, with neither any *ex-post* statistical intervention nor factor combinations.<sup>16</sup> The results presented in this article are important because they help to identify the most accurate group when forecasting inflation and, thus, help refine the information provided by the survey for inflation forecasting purposes.

---

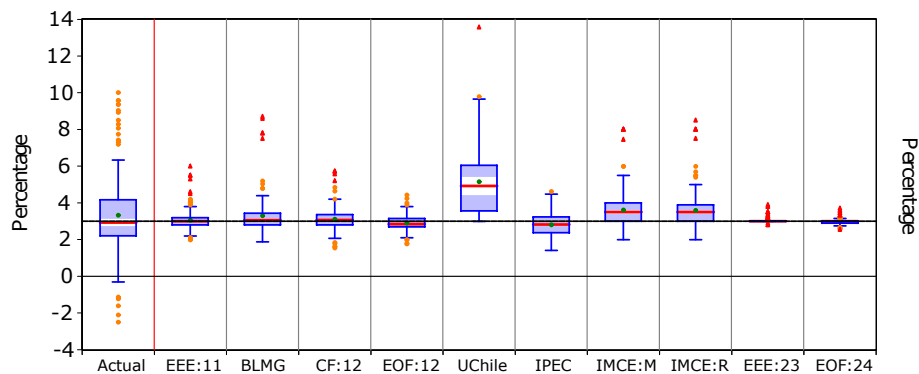
<sup>16</sup>This task, certainly important from a predictive point of view, could start by considering the methodologies proposed in [Clements and Hendry \(1996\)](#) and [Bentancor and Pincheira \(2010\)](#).



## Appendix A - Surveys of Inflation Expectations in Chile

In this appendix, I compare the Universidad de Chile's *Centro de Microdatos* inflation expectations average ("UChile") with seven surveys asking at the same horizon plus two at 24-months ahead, i.e., Chile's official monetary policy horizon, and including the actual total inflation series.

The comparison is made through the boxplots of [Figure A.1](#).<sup>17</sup> Note that "UChile" is the survey with the worse performance when using all raw data in terms of point and dispersion—with the distinctive feature that is the only survey considered in this figure that is conducted in a quarterly basis, whereas remaining ones are conducted in a monthly or daily basis. Other consumer expectations are "IPEC" ("*Índice de Percepción de la Economía*" elaborated by *Adimark*) whereas the remaining ones are answered by either experts or professional analysts. The best results at 12-month horizon are obtained with "EEE:11" ("*Encuesta de Expectativas Económicas*", elaborated by the Central Bank of Chile), "BLMG" (*Bloomberg* survey), "CF:12" (*Consensus Economics*



*Notes:* All data are transformed to annual variation in order to be compared to the 3% annual variation official inflation target. "Actual" stands for annual variation of total CPI inflation in monthly frequency. Sample: Jan-00/Jul-20 (247 observations). Source: National Statistics Institute. "EEE:11" and "EEE:23 plus "EOF:12" and "EOF:24" stands for "*Encuesta de Expectativas Económicas*" at 11- and 23-month horizon, and "*Encuesta de Operadores Financieros*" (prior to each Monetary Policy Meeting) at 12- and 24-month ahead; respectively. Sample: EEE: Sep-01/Jul-20 (227 observations); EOF: Dec-09/Jun-20 (127 observations). Source: Central Bank of Chile. "BLMG" stands for *Bloomberg* daily median inflation forecast, considering the last day of each month. Sample: Jan-08/Jul-20 (151 observations). Source: *Bloomberg*. "CF:12" corresponds to a 24-term weighted average between the "current year" and the "next year" horizons of the *Consensus Forecast* report to reflect a unique 12-month comparable horizon. Sample: Jan-00/Jul-20 (247 observations). Source: *Consensus Economics*. "UChile" stands for the Universidad de Chile's *Centro de Microdatos Survey of Perception and Expectations on the Economic Situation in Greater Santiago* quarterly data. Sample: 2005.II/2018.IV (55 observations). Source: *Centro de Microdatos*. "IPEC" stands for "*Índice de Percepción de la Economía*" (consumers) and the original index with 50 as a neutral value is extrapolated as 3% inflation rate. Sample: Mar-02/Jun-20 (220 observations). Source: *Adimark*. "IMCE" stands for "*Indicador Mensual de Confianza Empresarial*" (entrepreneurs) where "M" stands for "Manufacturing" and "R" for "Retail". Sample: May-05/Jun-20 (182 observations). Source: *Instituto Chileno de Administración Racional de Empresas (ICARE)* and *Universidad Adolfo Ibáñez (UAI)*.

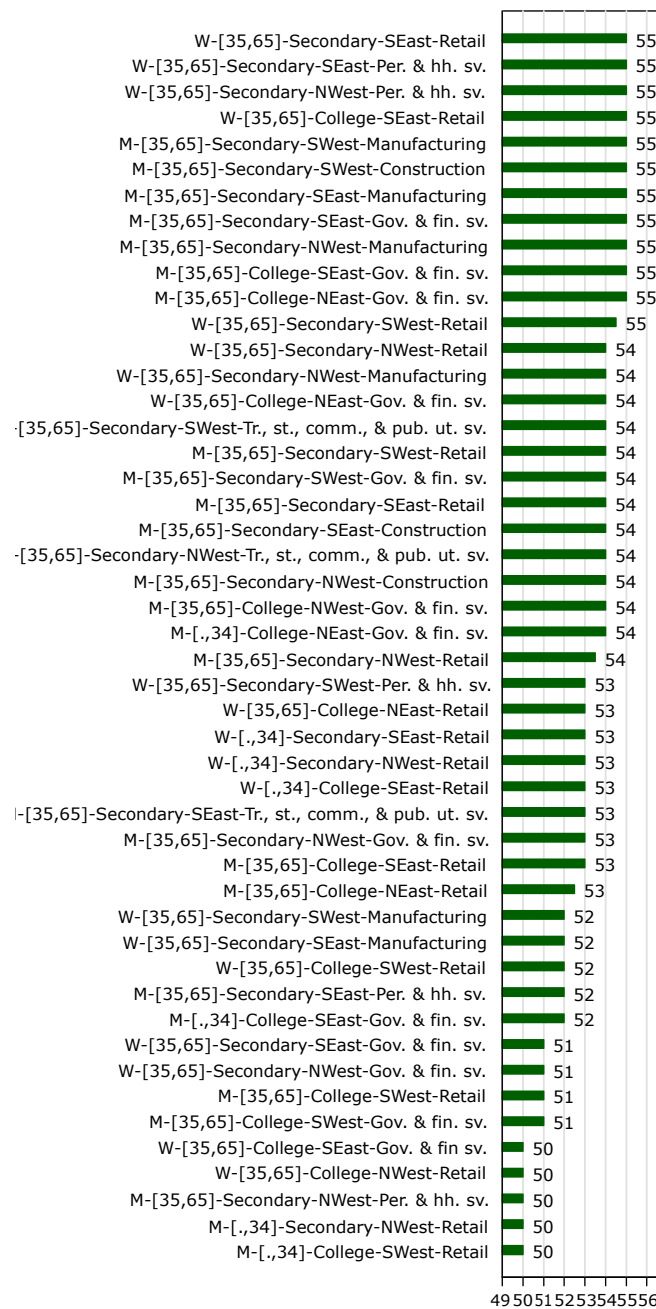
**Figure A.1.** Boxplots of Chilean inflation expectations surveys.

<sup>17</sup>Note that [Figure A.1](#) depicts survey-based inflation expectations only. However, it could be easily extended by making use of the information obtained from financial assets which are possible to extract the breakeven inflation rate. However, those expectations are of a different nature. The complete graph, however, including those expectations coming from financial assets, is available here: <https://drive.google.com/file/d/15qXSyOCncBg5SQ0RM-lJdBLW-87ZPZHF/view?usp=sharing>.

survey, applying a weighting scheme to transform it from a moving horizon to a 12-month fixed horizon), and “EOF:12” (*“Encuesta de Operadores Financieros”* conducted by the Central Bank of Chile before each Monetary Policy Meeting). The last two boxplots display the EEE and EOF inflation expectations at 24-month ahead, with a little or virtually no variation from the inflation target of 3% within the sample.

The “IMCE:M” and “IMCE:R” corresponds to the *“Indicador Mensual de Confianza Empresarial”* (entrepreneurs) elaborated by *Instituto Chileno de Administración Racional de Empresas* (ICARE) and *Universidad Adolfo Ibáñez* (UAI), where “M” stands for *“Manufacturing”* and “R” for *“Retail”*. These two surveys are not as accurate as those responded by experts and professional analysts, but certainly are much better than those of “UChile”. So, [Figure A.1](#) puts into perspective the challenge to be addressed and the sense in search for the best forecasters within the “UChile” employment survey.

## Appendix B - Considered Factors for Evaluation



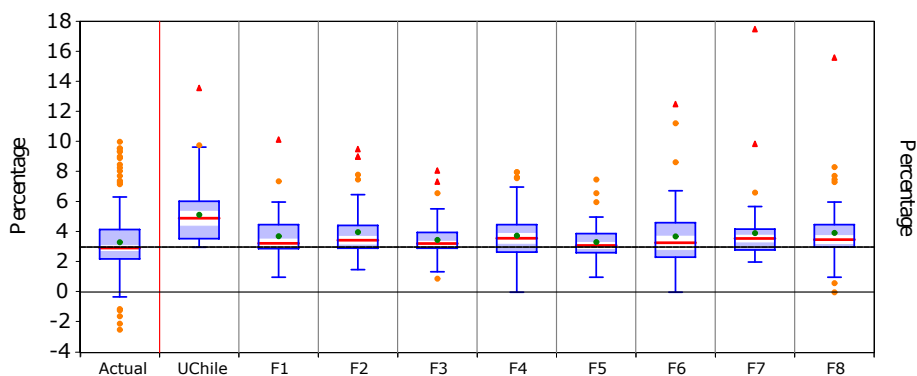
Notes: Factor key: [Gender]-[Age group]-[Education]-[Cardinal region]-[Economic sector]; see Table 1. Bars show the number of quarters with data considered in the analysis (those factors with at least 95% of total sample)—i.e., 55 out of 648 of total factors. Total sample: 2005.II-2018.IV (55 observations). Source: Author's calculations based on *Centro de Microdatos* database.

Figure B.1. Considered factors for evaluation.

### Appendix C - A Visual Comparison Between the Best Factors

In this appendix, I analyse the boxplots of the eight factors that outperform the RW (see Table 3) compared to the aggregate “UChile” factor and the actual total inflation series. The boxplots are shown in Figure C.1.

The first fact to note is that all factors display an enhanced forecast accuracy compared to the aggregate in terms of the mean point, whereas some of them (e.g.,  $F6$ ,  $F7$ , and  $F8$ ) are worst in terms of dispersion, i.e., a major number of outliers. Remarkably,  $F1$  display just a few outliers, similarly to  $F4$  and  $F5$ , but  $F4$  with a greater interquartile range. Also, the median of  $F1$ ,  $F3$ , and  $F5$  are the closest to the 3% target, but  $F3$  and  $F5$  display more outliers than  $F1$ , because they have a tighter interquartile range. Thus,  $F1$ ,  $F3$ , and  $F5$  factors comes out as the most promising factors within the survey and deserving more analyses delving into its differences.



Notes: All data show the CPI inflation annual variation. “Actual” stands for annual variation of total inflation in monthly frequency. Sample: Jan-00/Jul-20 (247 observations). “UChile” stands for the Universidad de Chile’s *Centro de Microdatos Survey of Perception and Expectations on the Economic Situation in Greater Santiago* quarterly data. Sample: 2005.II-2018.IV (55 observations). “F1” to “F8” stands for the inflation factors that outperform the RW—see Table 3  
 Source: Author’s calculations based on National Statistics Institute and Universidad de Chile’s *Centro de Microdatos* database.

Figure C.1. Boxplots of actual inflation and consumer-perceptions based inflation factors.

## Appendix D - RMSFE Ratio Results Using the Median Factor

**Table D.1**

RMSFE ratio and Giacomini-White test.

Factor	End-of-period	Average
M-[.,34]-NWest-Retail	3.801*	3.913*
M-[.,34]-SEast-Gov. & fin. sv.	2.745	2.859
M-[.,34]-NEast-Gov. & fin. sv.	1.081	1.090
M-[35,65]-SWest-Manufacturing	0.910	0.913
M-[35,65]-SWest-Construction	0.975	1.023
M-[35,65]-SWest-Retail	1.992	2.002
M-[35,65]-SWest-Gov. & fin. sv.	1.959	1.996
M-[35,65]-SWest-Tr., st., comm., & pub. ut. sv.	2.164	2.261
M-[35,65]-SEast-Manufacturing	0.852	0.872
M-[35,65]-SEast-Construction	1.072	1.107
M-[35,65]-SEast-Retail	1.570	1.567
M-[35,65]-SEast-Gov. & fin. sv.	1.129	1.186
M-[35,65]-SEast-Per. & hh. sv.	2.736	2.770
M-[35,65]-SEast-Tr., st., comm., & pub. ut. sv.	1.810	1.854
M-[35,65]-NWest-Manufacturing	0.930	0.934
M-[35,65]-NWest-Construction	1.573	1.593
M-[35,65]-NWest-Retail	0.998	1.017
M-[35,65]-NWest-Gov. & fin. sv.	2.162	2.198
M-[35,65]-NWest-Retail	1.112	1.123
M-[35,65]-NWest-Tr., st., comm., & pub. ut. sv.	0.857	0.838
M-[35,65]-SWest-Retail	1.154	1.193
M-[35,65]-SWest-Gov. & fin. sv.	0.978	0.982
M-[35,65]-SWest-Retail	1.087	1.091
M-[35,65]-SEast-Retail	0.914	0.916
M-[35,65]-SEast-Gov. & fin. sv.	0.822	0.833
M-[35,65]-SEast-Retail	0.920	0.935
M-[35,65]-NEast-Retail	1.958	1.984
M-[35,65]-NEast-Gov. & fin. sv.	0.866	0.862
M-[35,65]-NEast-Retail	0.921	0.910
M-[35,65]-NWest-Gov. & fin. sv.	0.898	0.890
W-[.,34]-SEast-Retail	1.978***	2.055***
W-[.,34]-NWest-Retail	1.304	1.319
W-[.,34]-SEast-Retail	2.037	2.063
W-[35,65]-SWest-Manufacturing	1.872*	1.921*
W-[35,65]-SWest-Retail	1.339	1.360
W-[35,65]-SWest-Per. & hh. sv.	2.324**	2.397**
W-[35,65]-SWest-Retail	1.064	1.073
W-[35,65]-SEast-Manufacturing	2.442	2.485
W-[35,65]-SEast-Retail	1.594	1.629
W-[35,65]-SEast-Gov. & fin. sv.	2.367	2.425
W-[35,65]-SEast-Per. & hh. sv.	1.854	1.861
W-[35,65]-SEast-Retail	2.200	2.238
W-[35,65]-NWest-Manufacturing	1.918	1.957
W-[35,65]-NWest-Retail	1.490	1.518
W-[35,65]-NWest-Gov. & fin. sv.	1.312	1.317
W-[35,65]-NWest-Per. & hh. sv.	1.110	1.111
W-[35,65]-NWest-Retail	0.939	0.957
W-[35,65]-SWest-Retail	1.441	1.475
W-[35,65]-SEast-Retail	1.733	1.768
W-[35,65]-NEast-Gov. & fin. sv.	1.005	0.992
W-[35,65]-NEast-Retail	0.924	0.922

Notes: See notes to Table 2. Source: Author's calculations.

## References

- Ahlburg, D. A. (1984). Forecast Evaluation and Improvement Using Theil's Decomposition. *Journal of Forecasting* 3(3), 345-351.
- Alvarado, J. M., and Medel, C. A. (2020). Granular Inflation Facts: Evidence from Chile. Central Bank of Chile Mimeo.
- Arias, A., and Kirchner, M. (2019). Shifting Inflation Expectations and Monetary Policy. Central Bank of Chile Working Paper No. 829.
- Armantier, O., Topa, G., van der Klaauw, W., and Zafar, B. (2017). An Overview of the Survey of Consumer Expectations. *FEBNY Economic Policy Review* 23(2), 51-72.
- Axelrod, S., Lebow, D., and Peneva, E. (2018). Perceptions and Expectation of Inflation by US Households. Finance and Economic Discussion Series No. 2018-073. Board of Governors of the Federal Reserve System.
- Bentancor, A., and Pincheira, P. M. (2010). Predicción de Errores de Proyección de Inflación en Chile. *El Trimestre Económico* LXXVII(1), 129-154.
- Box, G. E. P., and Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- Bryan, M. F., and Venkatu, G. (2001a). The Demographics of Inflation Opinion Surveys. Federal Reserve Bank of Cleveland *Economic Commentary* 2022-10, 1-4.
- Bryan, M. F., and Venkatu, G. (2001b). The Curiously Different Inflation Perspectives of Men and Women. Federal Reserve Bank of Cleveland *Economic Commentary* 2022-11, 1-4.
- Campelo, A., Malgarini, M., Bittencourt, V. S., and Velho, V. V. (2019). Inflation expectations of Brazilian consumers: an analysis based on the FGV survey. *International Review of Applied Economics* 33(4), 505-522.
- Central Bank of Chile. (2018). *Mercado Laboral: Hechos Estilizados e Implicancias Macroeconómicas*. Santiago de Chile: Central Bank of Chile.
- Central Bank of Chile. (2020). *Dinámicas y Determinantes de la Inflación en Chile*. Santiago de Chile: Central Bank of Chile.
- Centro de Microdatos. (2009). *Encuesta de Ocupación y Desocupación en el Gran Santiago: Manual de Codificación Actividad Económica [Survey of Employment and Unemployment in Greater Santiago: Economic Activity Coding Manual]*. Santiago de Chile: Centro de Microdatos, Universidad de Chile.
- Centro de Microdatos. (2016). *Diseño Muestral Encuesta de Ocupación y Desocupación en el Gran Santiago [Sample Design of the Survey of Employment and Unemployment in Greater Santiago]*. Santiago de Chile: Centro de Microdatos, Universidad de Chile.
- Chanut, N., Marcel, M., and Medel, C. A. (2019). Can Economic Perception Surveys Improve Macroeconomic Forecasting in Chile? *Economía Chilena* 22(3), 34-97.
- CLAPES-UC. (2020). *Metodología para la Construcción del IPC-Q1 o IPC de los Pobres*. Santiago de Chile: Centro Latinoamericano de Políticas Económicas y Sociales, Pontificia Universidad Católica de Chile.
- Clarida, R. H. (2012). What Has—and Has Not—Been Learned about Monetary Policy in a Low-

- Inflation Environment? A Review of the 2000s. *Journal of Money, Credit, and Banking* 44(S1), 123-140.
- Clements, M. P., and Hendry, D. F. (1996). Intercept Corrections and Structural Change. *Journal of Applied Econometrics* 11(5), 475-494.
- Coibion, O., and Gorodnichenko, Y. (2015). Is the Phillips Curve Alive and Well after All? Inflation Expectations and the Missing Disinflation. *American Economic Journal: Macroeconomics* 7(1), 197-232.
- Coibion, O., Gorodnichenko, Y., Kumar, S., and Pedemonte, M. (2020). Inflation expectations as a policy tool? *Journal of International Economics* 124, 103297.
- Diamond, J., Watanabe, K., and Watanabe, T. (2020). The Formation of Consumer Inflation Expectations: New Evidence from Japan's Deflation Experience. *International Economic Review* 61(1), 241-281.
- De Pooter, M., Robitaille, P., Walker, I., and Zdinak, M. (2014). Are Long-Term Inflation Expectations Well Anchored in Brazil, Chile, and Mexico? *International Journal of Central Banking* 10(2), 337-400.
- Drakos, K., Konstantinou, P. Th., and Thoma, F.-A. (2020). Inflation uncertainty and inflation expectations: Micro-level evidence from the eurozone. *The Journal of Economic Asymmetries* 21, e00150.
- Duca-Radu, I., Kenny, G., and Reuter, A. (2021). Can consumers' inflation expectations help stabilise the economy? European Central Bank Research Bulletin No. 79.
- Easaw, J., Golinelli, R., and Malgarini, M. (2013). What determines households inflation expectations? Theory and evidence from a household survey. *European Economic Review* 61, 1-13.
- Ehrmann, M., Pfajfar, D., and Santoro, E. (2017). Consumers' Attitudes and Their Inflation Expectations. *International Journal of Central Banking* 13(1), 225-259.
- Faure, F., and Medel, C. A. (2020). Does the Exposure to the Business Cycle Improve Consumer Perceptions for Forecasting? Microdata Evidence from Chile. Central Bank of Chile Working Paper No. 888.
- Galí, J., and Gertler, M. (1999). Inflation Dynamics: A Structural Econometric Analysis. *Journal of Monetary Economics* 44(2), 195-222.
- Ghysels, E., Osborn, D., and Rodrigues, P. M. (2006). Forecasting Seasonal Time Series. In G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting, Volume 1*. Amsterdam: North Holland.
- Giacomini, R., and White, H. (2006). Tests of Conditional Predictive Ability. *Econometrica* 74(6), 1545-1578.
- Goldfayn-Frank, O., Kocharkov, G. and, Weber, M. (2020). Households' Expectations and Unintended Consequences of Policy Announcements. Deutsche Bundesbank Research Brief 34th edition, November 2020.
- Gosselin, M.-A., and Khan, M. (2015). A Survey of Consumer Expectations for Canada. *Bank of Canada Review Autumn 2015*, 14-23.
- Goyal, A., and Parab, P. M. (2019a). Modeling Consumers' Confidence and Inflation Expecta-

- tions. *Economics Bulletin* 39(3), 1817-1832.
- Goyal, A., and Parab, P. M. (2019b). Modeling Heterogeneity and Rationality of Inflation Expectations across Indian Households. Indira Gandhi Institute of Development Research Working Paper No. 2019-002.
- Gürkaynak, R. S., Levin, A. T., Marder, A. N., and Swanson, E. T. (2007). Inflation Targeting and the Anchoring of Inflation Expectations in the Western Hemisphere. Federal Reserve Bank of San Francisco *Economic Review* 2007, 25-47.
- Henzel, S., and Wollmershäuser, T. (2008). The New Keynesian Phillips curve and the role of expectations: Evidence from the CESifo World Economic Survey. *Economic Modelling* 25(5), 811-832.
- Ichiue, H., and Nishiguchi, S. (2015). Inflation Expectations and Consumer Spending at the Zero Bound: Micro Evidence. *Economic Inquiry* 53(2), 1086-1107.
- Jean-Baptiste, F. (2012). Forecasting with the New Keynesian Phillips curve: Evidence from survey data. *Economics Letters* 117(3), 811-813.
- Jonung, L. (1981). Perceived and Expected Rates of Inflation in Sweden. *American Economic Review* 71(5), 961-968.
- Kenny, G., Kostka, T., and Masera, F. (2014). How Informative are the Subjective Density Forecasts of Macroeconomists? *Journal of Forecasting* 33(3), 163-185.
- Kenny, G., Kostka, T., and Masera, F. (2015). Density characteristics and density forecast performance: a panel analysis. *Empirical Economics* 48(3), 1203-1231.
- Larraín, M. (2007). Inflation Compensation and Inflation Expectations in Chile. *Monetaria* XXX(3), 305-329.
- Lindén, S. (2005). Quantified Perceived and Expected Inflation in the Euro Area-How Incentives Improve Consumers Inflation Forecasts. Paper presented at the OECD Workshop on International Development of Business and Consumer Tendency Surveys, Brussels.
- Łyziak, T. (2013). Formation of Inflation Expectations by Different Economic Agents. The Case of Poland. *Eastern European Economics* 51(6), 5-33.
- Łyziak, T., and Sheng, X. S. (2018). Disagreement in consumer inflation expectations. Narodowy Bank Polski Working Paper No. 278.
- Madeira, C., and Zafar, B. (2015). Heterogeneous Inflation Expectations and Learning. *Journal of Money, Credit and Banking* 47(5), 867-896.
- Malgarini, M. (2009). Quantitative Inflation Perceptions and Expectations of Italian Consumers. *Giornale degli Economisti e Annali di Economia* 68(1), 53-80.
- Marcel, M., Medel, C. A., and Mena, J. (2017). Determinantes de la Inflación de Servicios en Chile. Central Bank of Chile Working Paper No. 803.
- Medel, C. A. (2015). Inflation Dynamics and the Hybrid New Keynesian Phillips Curve: The Case of Chile. *Monetaria* III(1), 25-69.
- Medel, C. A. (2017). Forecasting Chilean Inflation with the Hybrid New Keynesian Phillips Curve: Globalisation, Combination, and Accuracy. *Economía Chilena* 20(3), 4-50.
- Medel, C. A. (2018). An Econometric Analysis on Survey-data-based Anchoring of Inflation Expectations in Chile. *Economía Chilena* 21(2), 128-152.



- Medel, C. A., and Pincheira, P. M. (2016). 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.
- Meyer, B., and Venkatu, G. (2011). Demographic Differences in Inflation Expectations: What Do They Really Mean? Federal Reserve Bank of Cleveland *Economic Commentary* 2011-07, 1-6.
- Mishkin, F. S. (2014). *Macroeconomics. Policy and Practice*, Second Edition. New Jersey: Pearson Education.
- Nason, J. M., and Smith, G. W. (2008). Identifying the New Keynesian Phillips Curve. *Journal of Applied Econometrics* 23(5), 525-551.
- National Statistics Institute (INE). (2015). *Encuesta Nacional sobre Uso del Tiempo, Síntesis de Resultados, La Dimensión Personal del Tiempo [National Survey on Time Use, Summary of Results, The Personal Dimension of Time]*. Santiago de Chile: National Statistics Institute.
- Newey, W. K., and West, K. D. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3), 703-708.
- Palmqvist, S., and Strömberg, L. (2004). Households' inflation opinion - a tale of two surveys. Sveriges Riskbank *Economic Review* 4, 23-42.
- Paloviita, M., and Mayes, D. (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.
- Paloviita, M. (2009). Estimating open economy Phillips curves for the euro area with directly measured expectations. *New Zealand Economic Papers* 43(3), 233-254.
- Pfajfar, D., and Santoro, E. (2008). Asymmetries in Inflation Expectations Formation Across Demographic Groups. Cambridge Working Papers in Economics No. 0824.
- Pincheira, P. M., and Medel, C. A. (2016). Forecasting with a Random Walk. *Czech Journal of Economics and Finance* 66(6), 539-564.
- Reis, R. (2006a). Inattentive Consumers. *Journal of Monetary Economics* 53(8), 1761-1800.
- Reis, R. (2006b). Inattentive Producers. *Review of Economic Studies* 73(3), 793-821.
- Rossouw, J., Padayachee, V., and Bosch, A. (2011). A Comparison of Inflation Expectations and Inflation Credibility in South Africa: Results from Survey Data. *South African Journal of Economic and Management Sciences* 14(3), 263-281.
- Sabrowski, H. (2008). Inflation Expectations Formation of German Consumers: Rational or Adaptive? University of Lüneburg Working Paper Series in Economics No. 100.
- Souleles, N. S. (2004). Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consume Sentiment Surveys. *Journal of Money, Credit and Banking* 36(1), 39-72.
- Theil, H. (1971). *Applied Economic Forecasting*. Amsterdam: North-Holland.