Economic Watch

Mexico

Consumer inflation expectations: determinants and threshold levels

April 5, 2013

Economic Analysis

Arnulfo Rodriguez
arnulfo.rodriguez@bbva..com

Pedro Uriz
pedro.uriz2@bbva.com

- **Consumer inflation expectations are largely optimistic[1], in line with the good average performance of headline inflation (3.8%) during the analyzed period.**
- **Consumer inflation expectations mainly react to the following components of the CPI: processed food, energy and public sector prices.**
- **Consumer inflation expectations turn pessimistic when the annual inflation of both processed food and the first lag of energy are at least 4.43% and 7.93%, respectively.**
- **By splitting the sample into women and men, both react to similar components of the CPI such as processed food and public prices; however, the second group does not show inflationary pessimism in the analyzed period.**

Measuring inflation expectations is of great importance for the monetary policy of central banks. In particular, measures which are reliable, frequent and timely become an essential input for monetary regimes based on inflation targeting schemes because, by construction, they are of a forward-looking nature. In regard to the latter, Berk and Hebbink (2010) identify three possible goals or uses for inflation expectations: (a) as a factor to be taken into account for determining the monetary policy stance that helps to reach the target; (b) as a tool to continuously monitor the credibility[2] by comparing them with the inflation target; and (c) as a valuable input to the inflation forecast process carried out by central banks.

Although inflation expectations are of key importance to central banks, their measurement is not an easy task. Lindén (2010) mentions that obtaining representative measurements for this type of expectations is an enormous challenge for economists and policy makers. The information of inflation expectations formed by various economic agents is obtained by extracting implicit expectations contained in financial instruments or by conducting surveys among professional forecasters or consumers. In order to better understand the inflation expectations of the general public, in this document we analyze the expectations of Mexican consumers.

Since one of the particular goals of this study is to find and sort the most relevant inflation components for the formation of those expectations, the Regression Trees methodology was used for mainly two reasons: (i) it is appropriate in view of both the relatively high number of inflation components which were tested and the relatively low number of observations in the estimation sample; and (ii) it also allows to sort the importance of components in the formation of inflation expectations, and to obtain the threshold levels above which consumers react to adjust the aforementioned expectations.

---

[1] Consumer inflation expectations are considered optimistic when expected inflation for the following 12 months is below current annual inflation.

[2] When the public does not have confidence in the central bank, its expectations will reflect on prices and wages decisions conducive to inflation levels above the policy target.

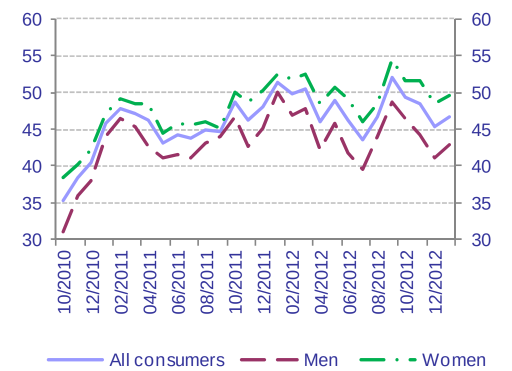
Data description and methodology

The information on Mexican consumers' inflation expectations was taken from the qualitative responses to question 12 of the basic questionnaire of the National Consumer Confidence Survey (Encuesta Nacional sobre Confianza del Consumidor, ENCO), which is formulated as follows:

Compared with the last 12 months, how do you think prices will perform in the country over the next 12 months?

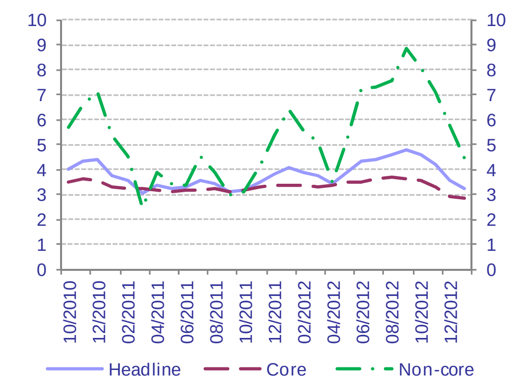
1. They will decrease a lot
2. They will decrease a little
3. They will remain unchanged
4. They will increase a little
5. They will increase the same
6. They will increase a lot
7. Do not know

Chart 1
Consumer inflation expectations (diffusion index)



Source: BBVA Research with INEGI data

Chart 2
Headline inflation and components (y/y % change)



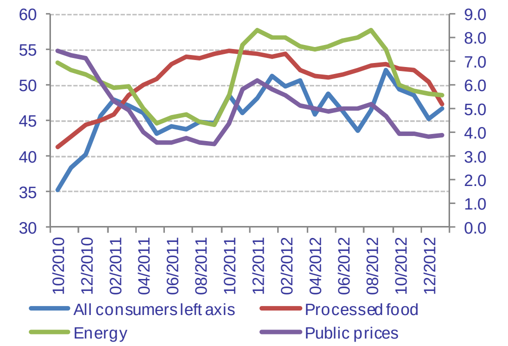
Source: BBVA Research with INEGI data

In order to have an aggregate measure of Mexican consumers' inflation expectations, a monthly diffusion index was constructed.[3] Data from October 2010 to January 2013 were used.[4] Diffusion indices by gender were also constructed in order to take into account the responses only by women or by men.

[3] The diffusion index was calculated as follows: (1) the "Do not know" response was prorated amongst the other response options to question 12; (2) the proportions of response to these options were calculated over total responses without considering expansion factors for households; (3) the weights of 0.0, 0.0, 0.0, 0.0, 0.5 and 1.0 were used for the "They will decrease a lot", "They will decrease a little", "They will remain unchanged", "They will increase a little", "They will increase the same" and "They will increase a lot" options, respectively; (4) each proportion was multiplied by its corresponding weight; and (5) the products of the last step were added together. It is worth mentioning that the allocation of the zero weight to the first 4 options is done so that the level of 50 will reflect neutral inflation expectations. In other words, an expectation of the same inflation for the next 12 months.

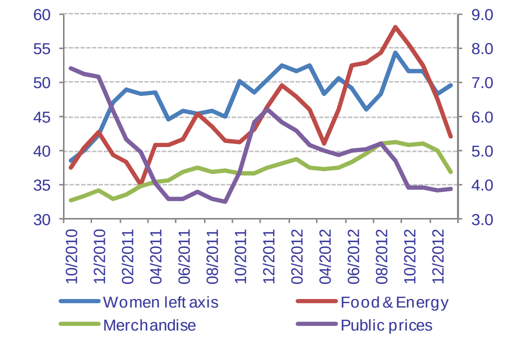
[4] The ENCO's microdata begin in January 2003. However, the diffusion index of inflation expectations was constructed from October 2010 given that the phrasing of question 12 was different in the months prior to that date.

Chart 3
Inflation expectations of all consumers and some inflation components (diffusion index and y/y % change)



Source: BBVA Research with INEGI data

Chart 4
Women's inflation expectations and some inflation components (diffusion index and y/y % change)



Source: BBVA Research with INEGI data

In the above charts, it is evident that pessimism periods (when the diffusion index is higher than 50) are relatively infrequent in the sample period. In other words, the pessimism threshold was only exceeded in 10.7%, 32.1% and 3.6% of the months for the indices of all consumers, women and men, respectively.

In regard to the explanatory factors that were explored, annual growth rates of the following indices were used (as well as their first two lags): general, core, goods, processed food, other goods, services, housing, education, other services, non-core, agricultural and livestock, agricultural, livestock, public (energy and public sector prices), energy, and public sector prices. Moreover, two *ad-hoc* indices were also constructed: one containing all the energy and food elements, and another containing all products other than energy and food.[5]

The Regression Trees methodology is part of the data mining estimation techniques.[6] In general terms, it could be described as a recursive partition of the response variable (or dependent) into several sub-groups, where each one of them is determined on the basis of the level of an explanatory factor (the variable selected in each partition can differ) from which the sum of squared errors of the observations found there is minimized.[7] It is important to note that the explanatory factors whose partitions are closest to the top of the tree are relatively more important to understand the studied phenomenon.

As to the question of how appropriate it might be to use the Regression Trees methodology, Fridley (2010) mentions that its use is recommended when there are compelling reasons to believe that the variables are not necessarily additive or when the number of explanatory factors to be tested is relatively large. Nevertheless, the same author warns that the trees tend to overfit the data and their topology is highly sensitive to small changes in explanatory factors.[8]

[5] The inclusion of an index that contains all energy and food prices obeys to both the fact that these products make up an important proportion of the consumers' basket (40% of total expenditures) and their high price sensitivity to local and global supply shocks.

[6] Data mining estimation techniques explain the relationship between explanatory factors and the variable of interest through the inductive method. In other words, they allow the data to define ex-post the variables which are part of the approximation function. In contrast, econometric estimation methods rely on the deductive method given that the generation of "knowledge" is based on general rules with the explanatory factors defined ex-ante.

[7] A mathematical explanation detailed on the Regression Trees methodology is presented in the appendix to this document.

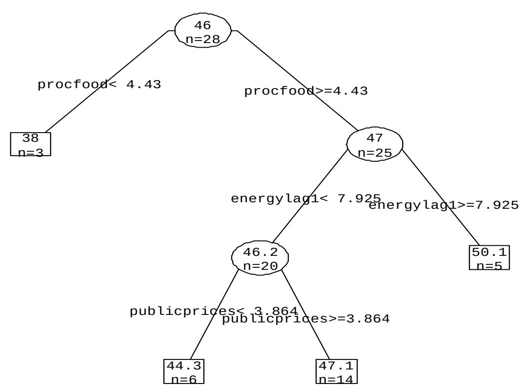
[8] In order to mitigate the overfitting problem, usually the tree is pruned using the cross validation procedure. In this document the aforementioned procedure is used and the results are only displayed for the pruned trees which were chosen for generating the minimum cross validation error.

In our opinion, Regression Trees are an adequate analytical tool for analyzing consumers' inflation expectations (for the next 12 months) for the following reasons: (1) only certain components of annual inflation are believed to be guiding the formation of consumers' inflation expectations aggregated through the diffusion index; (2) the purpose is to test a relatively high number of components of annual inflation with the aim of focusing on the type of goods and services whose price changes actually have the greatest influence on the movements of such expectations; and (3) the possible existence of threshold levels for some of the inflation components above which consumers probably react by changing their inflation expectations.

Analysis of results

The Regression Trees methodology was applied to the diffusion indices of all consumers, men and women.[9] For the first case, chart 5 shows that the most relevant components of annual inflation for forming expectations are: processed food, the first lag of energy and public prices (energy and public sector prices).[10] Furthermore, it can be inferred from chart 5 that in order for the diffusion index to exceed the barrier of 50.0 (or that inflation over the next 12 months will increase in relation to that of the previous 12 months), the annual inflation of processed food and the first lag of energy have to be at least 4.43% and 7.93%, respectively. For the case of women, it is evident from chart 6 that the most relevant factors are: the second lag of energy and food, public prices and the first lag of processed food. Their diffusion index shows inflationary pessimism when the level of these factors is higher than or equal to 4.59%, 3.81% and 6.67%, respectively. Finally, for the case of men, chart 7 shows that the most important components are: processed food and the first lag of public price inflation.

Chart 5
Regression tree for inflation expectations of all consumers

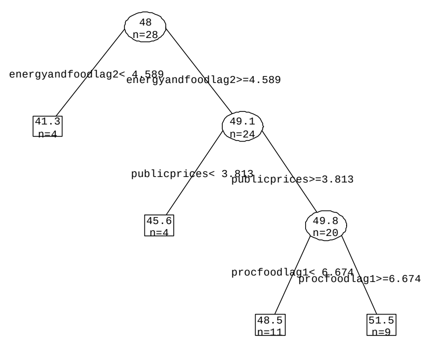


Source: BBVA Research with INEGI data

[9] The trees were generated with the rpart package, which is available to be used in the context of the R programming language.

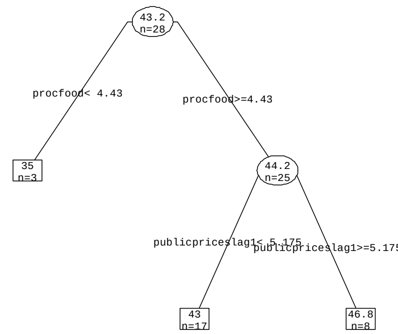
[10] In our opinion, these components turn out to be the most relevant due to three possible reasons: 1) the high purchase frequency of these goods and services; 2) their relative importance within total consumer expenditures; and 3) in the case of energy and food prices, their relatively high inflation during the analyzed period.

Chart 6
Regression tree for women's inflation expectations



Source: BBVA Research with INEGI data

Chart 7
Regression tree for men's inflation expectations



Source: BBVA Research with INEGI data

The comparison between the trees described in charts 6 and 7 poses some interesting questions. First, it is evident that men, unlike women, do not exhibit inflationary pessimism given that the tree shows that the average value of their index is always less than 50.0 in all terminal nodes.[11] Second, the food accompanied by energy or only processed food appear to be the main determinant underlying the formation of inflation expectations in both groups. This result could indicate that consumers might be reacting fundamentally to the evolution of food inflation because food purchases are relatively more frequent and, therefore, it is easier to identify supply shocks affecting food prices.

Conclusions

The most important findings of this analysis are the following: i) in general, inflation expectations of informants are predominantly optimistic, which is in line with the good performance of average headline inflation over the period (3.8%); ii) the diagrams of the regression trees show that consumers react (in descending order of importance) to the following inflation components: processed food, energy and public sector prices; iii) in order for the diffusion index of total consumers to show pessimism (index above 50), the annual inflation of processed food and the first lag of energy would need to be at least 4.43% and 7.93%, respectively, iv) for the case of women, the pessimism situation would occur when the level of the second lag of energy and food inflation (the ad-hoc constructed index), the inflation of public prices, and the first lag of processed food inflation were greater than or equal to 4.59%, 3.81% and 6.67%, respectively; v) finally, for the case of men, the tree does not indicate any path towards pessimistic expectations of inflation, with inflation of processed food and that of the first lag of public prices being the main variables to which they react.

[11] The apparent differences observed between unconditioned inflation expectations between women and men might largely be due to the fact that the first group has a lower average income than the second group, which would imply that the differences between genders were simply reflecting the difference in income levels. In support of this argument, Webley, Eiser and Spears (1988) find that a relatively high inflationary pessimism correlates with a relatively low economic optimism.

Appendix

This appendix describes the mathematical procedure used for generating regression trees shown in the text. For that purpose, the procedure of Hastie, Tibshirani and Friedman (2009) and the Cross Validation technique in Therneau and Atkinson (1997) are used as guides to grow regression trees and to prune them, respectively.

For a set of *s* explanatory factors $x_i = (x_{i1}, x_{i2}, x_{i3}, \ldots\ldots, x_{is})$ and a response variable (dependent) y_i for *i* = 1, 2, 3, 4,....., *N* observations, the following steps are followed to construct the trees:

1. Data are partitioned into *M* regions $R_1, R_2, R_3, \ldots\ldots, R_M$ and the response variable is modeled as a constant c_m in each region:

$$f(x) = \sum_{m=1}^{M} c_m I(x \in R_m)$$

where $I(x \in R_m) = 1$ if and only if $x \in R_m$

2. If the sum of squared errors is minimized $\sum_{i=1}^{N} (y_i - f(x_i))^2$ then the estimated value $\hat{c}_m$ will be equivalent to the average observations of y_i in the region R_m.
3. First, using all data, an explanatory factor is sought *j* with a threshold level *l* to separate the information into two regions: $R_1(j,l) = \{X \mid X_j \le l\}$ and $R_2(j,l) = \{X \mid X_j > l\}$.
4. The explanatory factor *j* and its threshold level *l* are found resolving the following optimization problem:

$$\min_{j,l}\left[\min_{c_1} \sum_{x_i \in R_1(j,l)} (y_i - c_1) + \min_{c_2} \sum_{x_1 \in R_2(j,l)} (y_i - c_2)\right]$$

5. For any pair of explanatory factor *j* and threshold level *l*, the estimated values $\hat{c}_1$ and $\hat{c}_2$ will be equivalent to the averages of the observations of y_i in R_1 and y_i in R_2, respectively.
6. Step 5 is carried out for all the explanatory factors to determine the best *j*, *l* pair.
7. Steps 3 to 6 are repeated for R_1 and R_2 and so on for all the resulting regions up until the minimum number of data per node is no longer met or until a further partition can no longer improve the model's fit by a given factor.

It is important to bear in mind that the trees shown in the text are pruned trees. The Cross Validation technique carried out to reach the sizes of trees which are less prone to the overfitting problem is outlined as follows:

1. The cost complexity function is defined as:

$$C_\alpha(T) = \sum_{m=1}^{|T|} N_m Q_m(T) + \alpha|T|$$

where

$\alpha \equiv$ parameter which penalizes the tree's growth

$m \equiv$ tree node

$|T| \equiv$ total number of terminal nodes

$N_m \equiv$ number of observations in the node

$Q_m(T) \equiv$ average sum of squared errors in the node

2. Given that it is possible to construct trees of the same size (with the same number of nodes) for intervals of α , there can be up to (z-1) number of trees assuming the following intervals: $(0,\alpha_1],(\alpha_1,\alpha_2],(\alpha_2,\alpha_3],......,(\alpha_{z-2},\alpha_{z-1}]$.
3. The sample is divided into s groups; all with the same number of observations.
4. One of the groups is left outside the estimation sample, and the (z-1) trees associated with the intervals of step 2 are constructed.
5. In order to form trees of a relative small size, the nodes which disappear from the next largest tree do so by having the smallest contribution to the increase in the sum of squared errors.
6. Using the data of the group which was not considered in the estimations in the previous step, the sum of squared errors is calculated in each one of the (z-1) trees.
7. Steps 4 and 5 are repeated for each one of the *s* groups;
8. For each one of the (z-1) trees, the ten calculations of the squared errors are added.
9. The tree with the lowest accumulated sum of squared errors is chosen.

References

Berk, J.M. and Hebbink, G. (2010). "The European consumer and monetary policy." In Sinclair, P. (editor). Inflation Expectations. Routledge, London and New York.

Fridley, J.D. (2010). "Tree models in R," Plant Ecology: Niche Modeling in R (BIO 793), class notes. Available at http://plantecology.syr.edu/fridley/bio793/cart.html

Hastie, T., Tibshirani, R. and Friedman, J. (2001). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second edition, 2009. Springer, New York.

Lindén, S. (2010). "400,000 observations on inflation perceptions and expectations in the EU: what do they tell us?" In Sinclair, P. (editor). Inflation Expectations. Routledge, London and New York.

Therneau, T.M. and Atkinson, E.J. (1997). "An Introduction to Recursive Partitioning Using the RPART Routines." Available at http://www.mayo.edu/hsr/techrpt/61.pdf

Webley, P., Eiser, R. and Spears, R. (1988). "Inflationary expectations and policy preferences," Economics Letters, 28-3, 239-241