FAO-OEA/CIE-IICA WORKING GROUP ON AGRICULTURAL AND LIVESTOCK STATISTICS FOR LATIN AMERICA AND THE CARIBBEAN

Twenty-sixth Session

Port of Spain, Trinidad and Tobago, 5-7 June 2013

Expanding agricultural price statistics, new derived indicators, data coverage and quality

Session 6: Recent advances in agricultural economic statistics

FAO Statistics Division
2013

The Statistics Division collects data on agricultural and food prices for a wide variety of commodities and products. FAOSTAT, its data dissemination platform, is one of the main data hubs for internationally comparable data on agricultural producer prices.

The recent food price crisis and its implications for food security underlined the importance for policy-makers to improve the monitoring of agricultural and food value-chains. With this objective in mind, the Price Team of the Statistics Division is currently working in 2 main directions: expanding the coverage of price statistics to measure prices from farm to fork, with an improved timeliness and increased frequency of data collection; and constructing a set of derived indicators to better capture and monitor price dynamics. These datasets and indicators will be consolidated and progressively disseminated in price profiles, which will provide a quick and easy access to the indicators and the underlying data, for different geographical scales.

This paper presents and describes the data currently collected on agricultural and food prices, as well as the indicators and price profiles that are currently under development.

1. Improving the measurement and monitoring of prices and price formation along the value-chain

a. Agricultural and food prices

Producer prices

FAOSTAT provides data on absolute producer prices for primary crop and livestock products to users through its Price Domain on an annual basis, and, starting from 2012, also on a monthly basis. Producer price indices (basic commodities and commodity aggregates) calculated on the basis of absolute prices are also disseminated on an annual basis.

The price domain of FAOSTAT has the largest coverage of producer prices in the world, with data covering 136 countries and about 200 commodities, representing roughly 97 percent of the world's value of agricultural production at 2004-2006 international dollar prices. Absolute producer prices are available in local currency, standard local currency and US dollars. To date, prices are available from 1991 to 2010, while 2011 update is in progress. Continuous efforts are made not only to expand country coverage but also to improve data quality.

Data are collected annually through a FAO questionnaire, with 100 to 110 responses (an increasing number of countries - 60 in 2012 - provide monthly prices). Producer prices refer to prices received by farmers, i.e. prices determined "at the farm gate" or at the first-point-of-sale when farmers participate in their capacity as sellers of their own products. While the aim is to remain as close as possible to this concept, flexibility has to be allowed for given local circumstances. Indeed, the choice of appropriate selling points may differ depending on the country and on the marketing structure for each commodity. For example, wholesale or even local market prices may be appropriate proxies of farm-gate prices when the marketing chain is very limited. On the other hand, in more intermediated markets in which transport and commercial margins constitute a significant share of the final product price, it is necessary to remain as close as possible to the farm-gate concept.

Quality and completeness of data varies considerably across countries. Although the general concept is to measure prices "at the farm-gate", many countries provide information at the first point of sale (wholesale markets in most cases). For monthly prices, this problem is even more acute: in some cases, prices provided are in fact closer to the concept of retail prices. Data quality checks are performed thanks to automatized routines and interactions with data providers to require explanations and revisions to the data when necessary.

Missing data are estimated based on country and commodity specific techniques (use of commodity aggregate averages, international commodity prices, time-series techniques, etc.). See Box 1 for details on estimation method. According to the current dissemination policy, only interpolations and retropolations are disseminated on FAOSTAT, while extrapolations are only used in the computation of derived indicators (value of agricultural production, price indices). Validation and estimation procedures have been the object of continuous revision aimed at improving data quality and consistency.

Annual Producer Price Indices are compiled and disseminated each year on the basis of official and estimated absolute producer prices, using as weighting variable the value of agricultural production.

Box 1 Imputation methods for missing prices

Missing information is estimated on the basis of:

Alternative semi-official or non-official data sources Examples are absolute producer prices disseminated by EUROSTAT for EU countries and the USDA/ERS for the USA.

Auxiliary information The estimation is carried out either by applying the growth rate of the auxiliary variable to the previous observed level of the missing price (or average/median growth rate if several variables are used) or by including these variables in linear regressions.

Official prices for the corresponding commodity group (e.g. average cereal price change to impute millet prices); Official price of the "nearest" commodity (e.g. orange prices to impute lime prices); wholesale or retail prices, such as those disseminated on the GIEWS platform; International market prices for agricultural commodities (FAO price indices, World Bank Pink Sheets, etc.); unit values; and aggregated price indices and deflators (agricultural producer price indices, food consumer price indices, agricultural GDP deflator, etc.).

Interpolations and trend regressions Linear interpolation may be used when there is a gap between two official data points and the linearity assumption is plausible, at least locally. This technique is only used when the gap to be filled is not too wide, i.e. not more than 3-4 missing data points; Trend regressions can be used either for filling data gaps or for extrapolations when a linear/exponential/polynomial trend is plausible.

Price ratios based on Technical Conversion Factors (TCFs) This method is applicable to impute missing data for processed commodities, in which a set of pre-established extraction rates are used to calculate the quantity of a processed product that can be obtained from a given amount of primary commodity, or vice versa (for example for livestock, oil crops, or some fibre crops).

Consumer prices

With a view to extend the statistical coverage of agricultural and food value-chains, the Statistics Division has started in 2011 a collaboration with the International Labour Organization to collect country data and metadata on monthly consumer price indices (Food and All Items index). Annual and monthly CPIs are published for some 130 countries[1] and updated each month.

Data timeliness, completeness and consistency varies considerably across countries. For example, some countries include food and beverages in the Food index while others only account for food products; some countries include food bought away from home while others only track food purchased by households in retail stores, etc. The timeliness of the data varies also considerably because of different reporting practices of countries and because of the time spent in data compilation by the ILO. While countries are required to provide indices based 100 in 2000, many countries use different base years. More problematically, many of the time series are provided with different base years over the period and are not linked together, leading to breaks in the series.

b. Analysis of price dynamics and relationships

Price levels, trends and variability reflect supply and demand conditions in agricultural and food markets. Indicators adequately measuring these three dimensions are needed in order to provide policy makers and analysts with appropriate measures on the basis of which forecasts can be made and policies assessed.

Price trends

Fixed-base indices, compiled and disseminated on FAOSTAT for both producer and consumer prices, measure the annual change in prices compared to a given base year (generally 2000). Price trends can be compared by computing the change in the Food Consumer Price Index relative to the Total Agricultural Producer Price Index. Significant and persistent differences in the respective trends of these indices may reflect changes in production and distribution margins, caused by modifications in the structure of the market (e.g. higher or lower concentration).

At the moment, comparisons can only be carried out on a country-by-country basis. However, as regionally aggregated indices will be available in the near future, regional as well as global-level comparisons might be undertaken (see 2.c for details on regional aggregations).

Volatility indicators

The level of price volatility or variability in agricultural and food markets is a good indicator of prevailing market conditions. At the upstream/producer level, a high level of price

[1] Some countries provide indices for urban or local areas only and a very limited number provide categorical disaggregation (low income vs. high income households, etc)

volatility may indicate that commodity supply is insufficient to cover demand and/or that producers are exposed to price fluctuations in agricultural inputs (e.g. fuel, feed, etc.) which are transmitted to output prices. At the downstream/retail level, price volatility may be caused by a high rate of transmission of producer or international commodity prices to the retail level, possibly reflecting short value-chains.

Appropriately measuring the level of price volatility and underlying market uncertainty is also important as market actors tend to modify their behaviour in presence of uncertainty: in an environment characterised by high price volatility, producers are less likely to invest given the uncertainty in their expected revenues; in the absence of clear price signals, consumers will less easily and quickly adapt their consumption behaviour; the effectiveness of policy interventions in fields such as price support policies, strategic stocks, regulation of commodity derivatives markets, etc. is determined to some extent by the perceived uncertainty in agricultural and food markets, of which price volatility is one of the measures.

The Price Team has been recently working on the development of a range of indicators measuring the amplitude of changes in consumer prices. These indicators, currently in their testing and validation phase, will be compiled and updated on a monthly basis. They will first be available on a country by country basis and then extended to higher geographical scales. A detailed description of the volatility indices is provided in Box 2 and a full example in Annex 2 .

The same methodology will be applied to monthly producer prices when these will be available for a sufficiently long time period. The comparison of price volatility at different levels of the value-chain is indicative of the extent to which price changes are transmitted along the chain. Producer price volatility is likely to be significantly higher than the volatility in food consumer prices as producers are directly exposed to a variety of shocks directly affecting their gross margins: shocks on markets for inputs, changes in international commodity prices and weather-related events affecting production. The extent to which the resulting variability in producer prices is transmitted to food consumer prices essentially depends on the length of the value-chain, on the market power of each actors of the chain, on the nature of the demand for that commodity or product and on the existence of possible substitutes.

Box 2 Consumer food price volatility indicators

Indicators should be able to capture the magnitude consumer prices change overtime, giving equal weights to increases and decreases. These indicators can be based on measures of observed volatility (i.e. standard statistical measures of dispersion) or on the modeling of the volatility process.

1. Volatility indicators based on measures of observed volatility

Statistical measures of dispersion on which volatility indices can be based include:

Standard deviation : $\sigma_{\mathrm{T}} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x}_T)^2}$, where x_t is the monthly growth rate of CPI or Food CPI and $\bar{x}_T$ is the average over $1 \dots T$.

This measure takes into account all the information over the period considered but is sensitive to extreme values. Extreme values or outliers can be meaningful and reflect actual market conditions but they can also result from inconsistencies in the data (change of base period, scope, etc.).

Range: $\mathrm{diff}_T = \max_{t=1\ldots T}(x_t) - \min_{t=1\ldots T}(x_t)$

This measure is by construction extremely sensitive to outliers. However, it can be meaningful as behavioral changes are often triggered by extreme variations and "psychological" thresholds.

Inter-quartile range: $\mathrm{iqr}_T = \mathrm{quantile}_{t=1\ldots T}(x_t, 75\%) - \mathrm{quantile}_{t=1\ldots T}(x_t, 25\%)$

Median of the absolute deviations from the median:

$$\mathrm{mad}_T = k.\mathrm{median}_{t=1\ldots T}\left(\left(x_t - \mathrm{median}_{t=1\ldots T}(x_t)\right)_t\right)$$

This measure is similar to the inter-quartile range but even less sensitive to extreme values. It converges towards the standard deviation as the sample size increases. The value of k depends on the underlying probability distribution of (x_t): assuming normal distribution, $k \cong 1.5$.

In order to capture the dynamics in food price changes overtime, the indicators proposed need to be based on a dynamic version of the dispersion measures listed above. This can be done by using sliding time periods, i.e. by choosing a fixed period length (number of months) and moving the time span as new data points become available.

The whole sequence of sliding standard deviations with time span h would be the following: $\boldsymbol{\sigma_h} = \left(\sigma_{[1,\ldots,h]}, \sigma_{[2,\ldots,h+1]}, \ldots, \sigma_{[t-h+1,\ldots,t]}, \ldots, \sigma_{[T-h+1,\ldots,T]}\right)$, where $\sigma_{[t-h+1,\ldots,t]}$, or more simply $\sigma_{t,h}$, is the standard deviation at time t conditional to the information generated by the last h data points.

Indices are then constructed by choosing an appropriate base period, for example the same than the one used for the CPI indices (currently base 100=2000).

2. Volatility indicators based on the modeling of the volatility process

The price of a commodity or a group of commodities at one point in time can be decomposed into its expected mean given the information available up to the preceding period and a random term. This random term represents the unexpected shocks that affect prices and, in the case of commodity prices in particular, these shocks are likely to be correlated overtime. These so-called GARCH processes (Generalized AutoRegressive Conditional Heteroscedasticity) can be reproduced by setting and estimating the structure of this autocorrelation. In its simplest version, this approach can be expressed mathematically by:

$$p_t = \bar{p}_t + \varepsilon_t \sigma_t \quad [1],$$

$$\sigma_t^2 = \alpha + \beta \varepsilon_{t-1}^2 \sigma_{t-1}^2 + \gamma \sigma_{t-1}^2 + \vartheta_t \quad [2].$$

where ε and ϑ are independently and identically distributed random terms and σ_t the conditional standard error.

The coefficients of equation [2] are generally determined by Maximum Likelihood Estimation (MLE), after setting initial values for the conditional variance. The conditional variance is then estimated iteratively over the whole period.

Price transmission indicators

The understanding of how and to what extent price changes are transmitted from international to national markets (horizontal transmission) and along agricultural and food value-chains (vertical transmission) is essential in assessing the exposition and vulnerability of market actors and consumers to price shocks. Measures of price transmission are needed to provide answers to questions such as: what is the overall impact of a price shock at the producer level on wholesale, retail and final consumer prices ? Is the shock fully transmitted or is it progressively absorbed and cushioned as it passes through the chain ? What is the speed of transmission ? Are national markets exposed to price shocks on international markets ?

Quantifying vertical price transmission consists in providing a measure for the size and the speed of the pass-through of a price shock (producer level) on consumer prices (retail level). For horizontal price transmission, the impact of a change in international prices on domestic prices (at wholesale or retail level) is investigated.

The Price Team of the Statistics Division is currently testing a framework to estimate price transmission coefficients on a reduced set of countries and commodities. The econometric approach, widely used in the literature, consists in estimating short and long-term price elasticities and to calculate impulse response functions. The later constitute appropriate measures of the size and timing of price transmission.

The theoretic as well as the econometric framework for estimating price transmission in agricultural and food markets is therefore relatively well grounded (cf. Annex 1 for more details on the econometric approach and an example). However, a series of practical challenges reduces the possibility to carry out estimations in a systematic way for a given group of countries, commodities and markets. Even when estimations may be undertaken, their significance might be weak and their interpretation delicate.

Data availability at the appropriate levels of the value-chain is one of the challenges that analysts have to face: to estimate vertical transmission, a sufficient number of price quotations have to be available at the producer, wholesale and retail levels for a similar commodity or product.

The choice of the market also influences the estimations: for horizontal transmission, a representative international price needs to be chosen for the commodity, which will likely differ from country to country according to its trade structure. It has also been argued that transmission is best estimated for specific markets at the local level (Benson and Faminow, 1985), but this information is seldom available.

Policy interventions, such as minimum or maximum purchasing prices, export or import restrictions, production subsidies, etc. and their evolution over time alter the degree of pass-through and may lead to weak transmission or inconclusive estimations.

Given these challenges, the estimation framework needs to be adapted to each country and market specificity.

c. Regional aggregations

Regional and world indices are useful for a wide range of reasons, among which: the possibility to compare price changes in one country to the regional or world average; to provide analysts monitoring food and agricultural markets with a global picture of price changes at world, regional and sub-regional levels; to carry out regional food security analysis.

The Statistics Division is currently finalising the development of regional consumer food price indices. These indicators, compiled on a monthly basis and on the basis of the regional disaggregation used by FAOSTAT, will complete the set of regional indices already existing in other data domains, such as production or trade.

Few organizations seem to compile and disseminate regional food price indices on a regular basis. The ILO's Global Trends provide indices with very limited regional disaggregation (developed and developing countries); the IMF provides a detailed regional disaggregation, but only disseminates an all items index; the regional scope of OECD's and Eurostat's indices is limited to their respective member countries. Furthermore, while all these organizations use GDP weights to aggregate country data, FAO uses weights based on country population, which is best adapted to keep focus on food security.

In a second step, regional consumer price indices will be completed by: regional producer price indices, calculated on a annual and monthly frequency on the basis of the data collected each year from country authorities, and; estimates of absolute producer prices at the regional and global levels, expressed in a single international currency (such as the international Dollar, currently used to carry out regional aggregations for the value of agricultural production).

2. Construction of price profiles

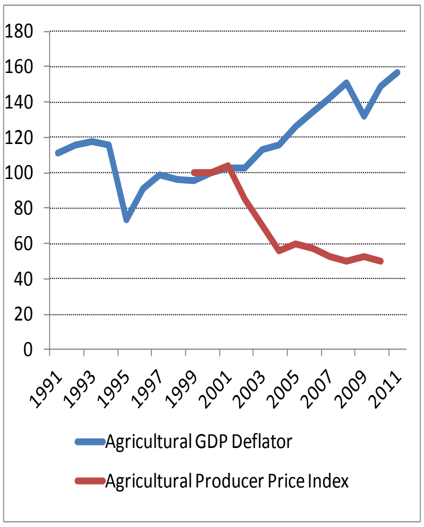
The primary data and derived indicators presented above will be progressively presented in price profiles, one for each level of regional disaggregation (country, regional and global levels). Examples of how these profiles will be structured are provided in this section.

a. Country profiles

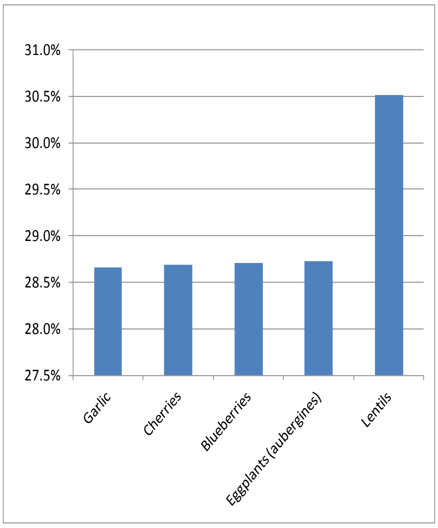
Country specific information on producer prices (absolute levels and indices), consumer prices (indices) and price transmission coefficients will be presented in profiles possibly structured in 3 parts - producer prices, consumer prices and value-chain indicators – as illustrated in the mock-ups below.

Producer prices

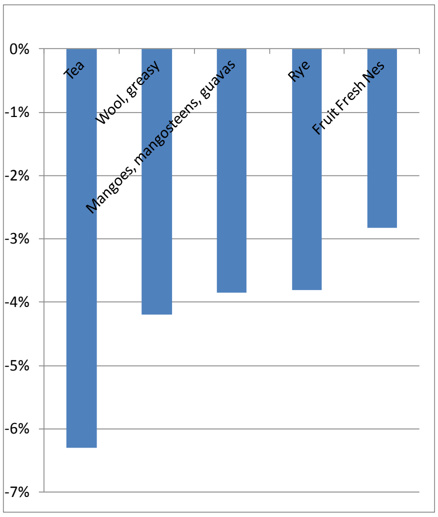
Price indicators



5 largest increases



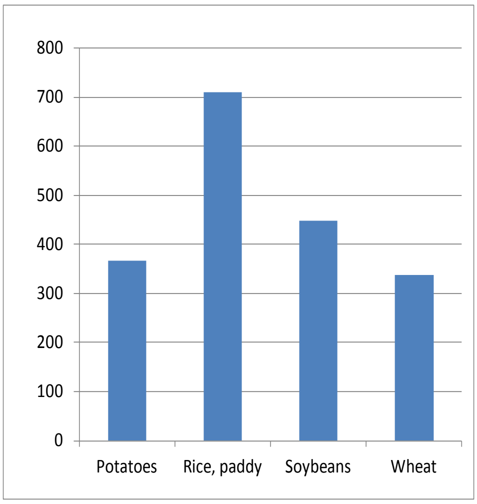
5 largest decreases



b. Regional profiles

Regional profiles will have a similar structure to country profiles but will be based on appropriate regional indicators. They will not include price transmission coefficients but may include average of price levels for a selected commodities. An example is provided below:

Price of the region's main commodities (avg. 2006-2010, USD/t)

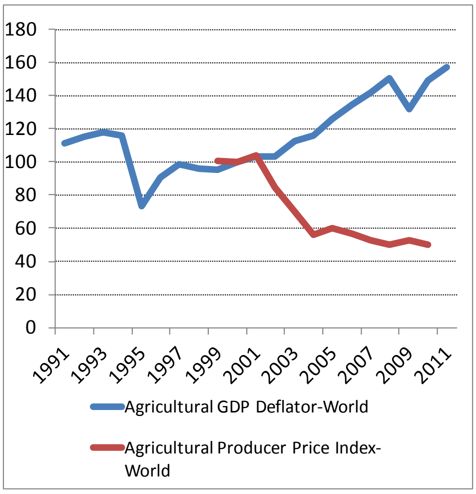


c. World profiles

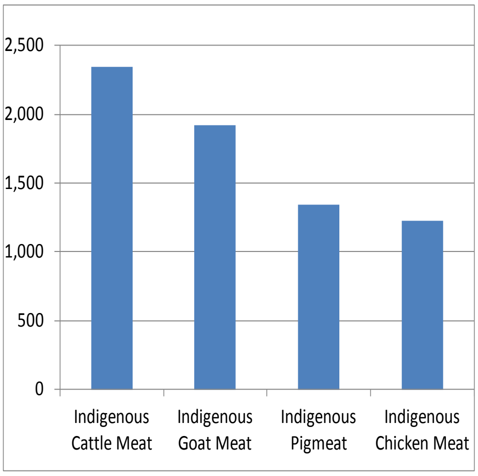
World profiles will display consumer and producer price indices computed at the global scale. Average producer prices for the main commodity groups might also be disseminated, either in USD or International Dollar terms, as illustrated in the examples below.

Producer Prices

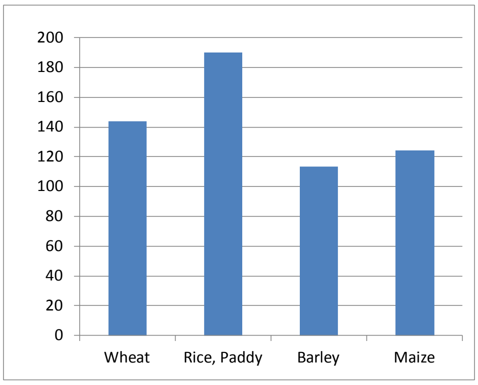
Price indicators



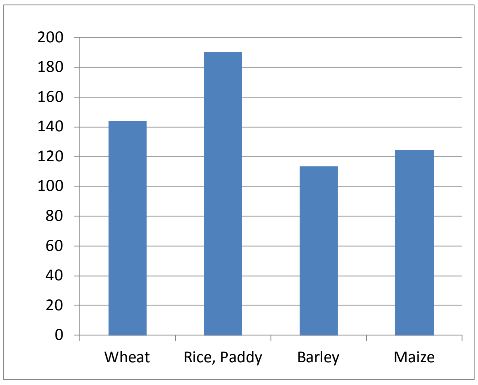
Livestock (avg. 2011 prices, I$/t)



Vegetables (avg. 2011 prices, I$/t)



Cereals (avg. 2011 prices, I$/t)



3. Points for discussion and requests

The FAO would like IICA member countries to provide their feedback, comments and inputs on these new lines of work. In particular, the Statistics Division would like to know if IICA members support the proposition to construct price profiles as well as the choice of new regional indicators, volatility measures and price transmission indicators.

In addition, in order to facilitate the up-coming data collection processes necessary to support this work, we would be thankful if member countries could provide the FAO with up-to-date contact details (organisation, name of contact officer) for the collection of prices at various points of the value-chain, by writing to price-statistics@fao.org.

ANNEXES

Annex 1 A possible framework for the estimation of price transmission

The impact of a given price shock at one point of the value-chain (or on international markets) on downstream prices (or domestic prices) can be estimated in an econometrically sound way using error correction models. These are designed to model non-stationary and co-integrated time series, as it is often the case for agricultural and food prices. Impulse response functions based on the estimated coefficients are then computed to measure the size and speed of price transmission.

1. Estimating price transmission coefficients

The objective is to measure the impact of a given change in the upstream price, P_t^U, of a given commodity, on the downstream price P_t^D.

Assuming that P_t^U and P_t^D are non-stationary time series and that there is a long-term or equilibrium relationship linking the two variables, the following error-correction model can be specified:

$$p_t^U = c_1 + \sum_{i=1}^{p} a_i p_{t-i}^U + \sum_{j=0}^{k} b_j p_{t-j}^D + r[P_{t-1}^U - (c_2 + bP_{t-1}^D)] + e_t$$

Where p_t^U (resp. p_t^D) is the first difference in upstream (resp. downstream) prices and e a stationary random variable with mean 0 and constant variance (error term).

The coefficients $(a_1, \dots, a_i, \dots, a_p)$ define the autoregressive structure of the model, $(b_0, \dots, b_j, \dots, b_k)$ are short-term price transmission coefficients, b the long-term price transmission coefficient and r the measure of the speed at which price levels go back to their equilibrium level after a shock.

Error correction models can be estimated using the two-step procedure suggested by Engle and Granger: first, consistent estimates of the long-term parameters are first determined by OLS regression; second, the lagged residuals of the long-term regression are inserted in the model in first differences and short-term parameters estimated by OLS or GLS regression.

Intuitively, if r is found to be not significantly different from 0, the error-correction mechanism is irrelevant and the time series would be better modeled by a simple augmented autoregressive model of the form:

$$p_t^U = c_1 + \sum_{i=1}^{p} a_i p_{t-i}^U + \sum_{j=0}^{k} b_j p_{t-j}^D + e_t$$

This model can be refined in a variety of ways by: adding exogenous explanatory variables in the error-correction model (measures of market concentration, policy variables, etc.); allowing for asymmetries in the reaction of downstream to upstream prices (e.g. different responses for price increases and decreases); allowing for threshold effects to account for different price regimes; assuming that the causality goes in both ways, from upstream to downstream prices and conversely, a bi-variate autoregressive or error correction model can be specified, where the first equation is the one presented above and the second one where p_t^D is endogenous and p_t^U is the explanatory variable.

Vavra P. and B. K. Goodwin (2005) provide more details on these models and their relevance to estimate price transmission in agricultural markets.

2. Impulse response functions : the measure of price transmission

Impulse Response Function (IRF) is the term used to define the impact of a shock in one of the explanatory variables on the endogenous variable of a dynamic system. In the case of linear models, the time profile of the impacts depends neither on the timing of the shock nor on its size. IRFs are themselves linear with respect to the size of the shock: the IRF relative to a 10% shock is equal to 2 times the IRF relative to a 5 % shock, etc.

In non-linear models, for example when asymmetries in price transmission and/or threshold effects are introduced, the IRFs are no longer linear and their size and time profile depends on past values (e.g. if a threshold is reached or not). Their interpretation is less straightforward than for linear IRFs.

3. Example

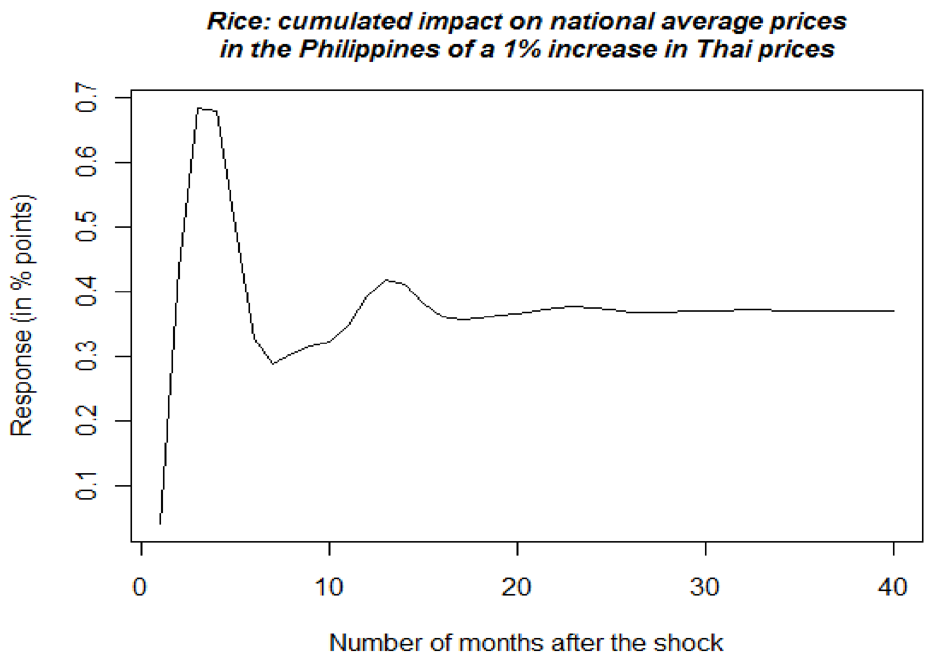
To illustrate the modelling approach described in 2., we propose to estimate the transmission of changes in international rice prices to wholesale rice prices in the Philippines (horizontal price transmission). Rice (25% broken) prices in Bangkok are taken as the international reference price (*Rice_th*). The national average price of rice (regular milled) is used for the Philippines (*Rice_ph*). Prices are taken in national currencies. The results are shown below:

Transmission equation: $\Delta Rice\_ph_t = a + \sum_{k=1}^{4} b_k \Delta Rice\_ph_{t-k} + \sum_{l=1}^{4} c_l \Delta Rice\_th_{t-l}$

Estimated transmission coefficients: $c_1 = 0.04, c_2 = 0.37, c_3 = -0.03, c_4 = -0.11$

These coefficients alone do not say much on the extent and timing of the transmission. In the absence of a co-integrating relationship, as it is the case in this example, it is difficult to infer from the estimated relationship the extent of the long-term pass-through of international prices to national prices. IRFs are needed to estimate the extent of price transmission and its distribution overtime.

Applied to our example, the IRF provides a measure of the impact of a given change of $Rice\_th$ (+1%, +10%, one standard deviation, etc.) in t on $Rice\_ph$ in $t+1, t+2, \dots, t+h, \dots, t+H$, everything else being held equal.

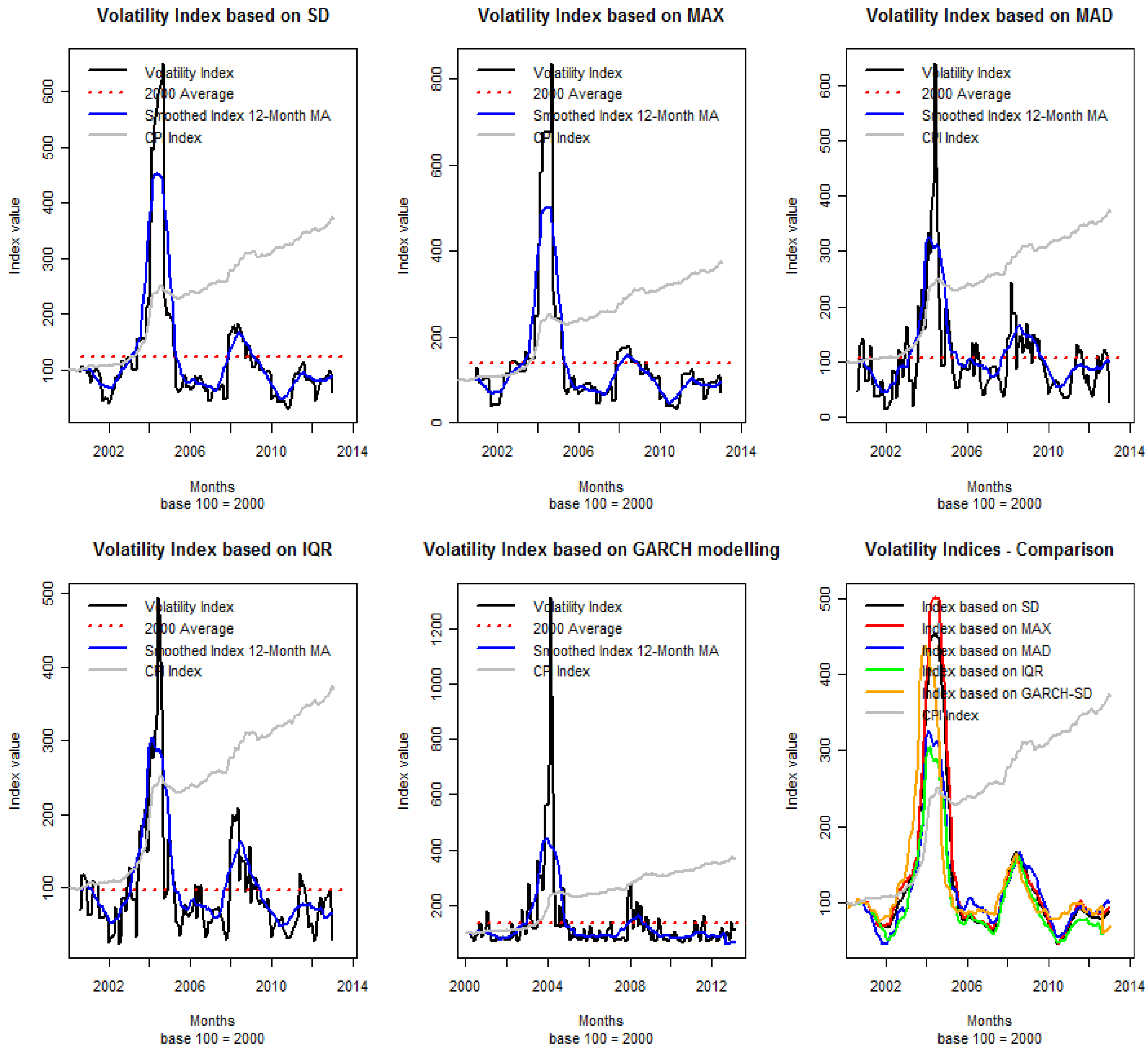


Extent of transmission: a 1% increase in Thai rice prices results in a long-term cumulated rise of 0.4% of rice prices in the Philippines.

Speed of transmission: the transmission is spread approximately over 24 months (the IRF converges towards 0.4 quickly after the 18th month). The shock is felt from the 2nd month and the impact inflates up to the 5th month, before receding between the 6th and the 9th month (cf. negative coefficients for the 3rd and 4th lag in the transmission equation) and stabilizing around 0.4 after that.

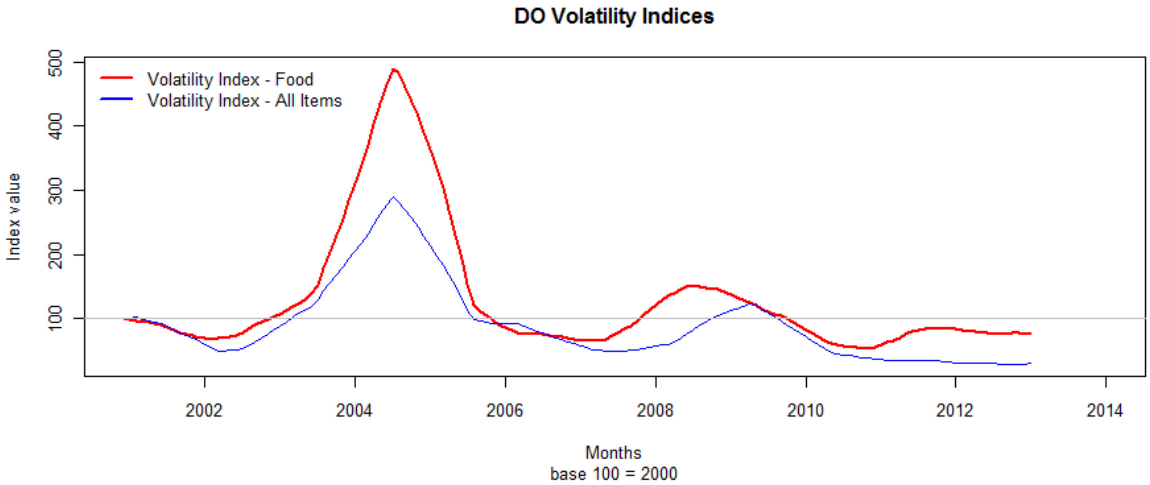
Annex 2 Computing volatility measures: the example of the Dominican Republic

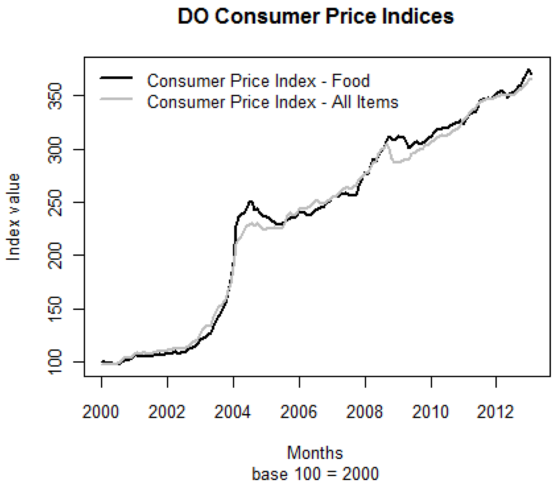
1. Comparison of different volatility measures

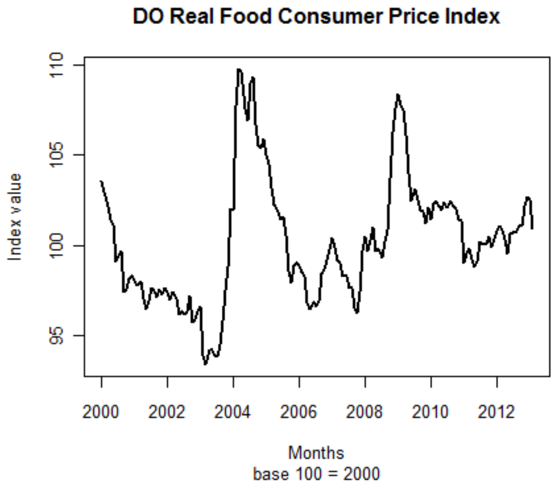


Notes: SD stands for standard-deviation based volatility measure, MAX for the difference between maximum and minimum values, MAD for the median of the absolute deviation from the median, IQR for the inter-quartile range and GARCH for the model-generated volatility.

2. Compared volatility in Food CPI and All items CPI







Note: Volatility measure based on the standard deviation

3. Comparison of Food volatility in the Dominican Republic with countries of the region (base 100=2007)