The Hierarchical Structure of Price Changes and Core Inflation

Paulo Picchetti

Instituto Brasileiro de Economia - Fundação Getulio Vargas

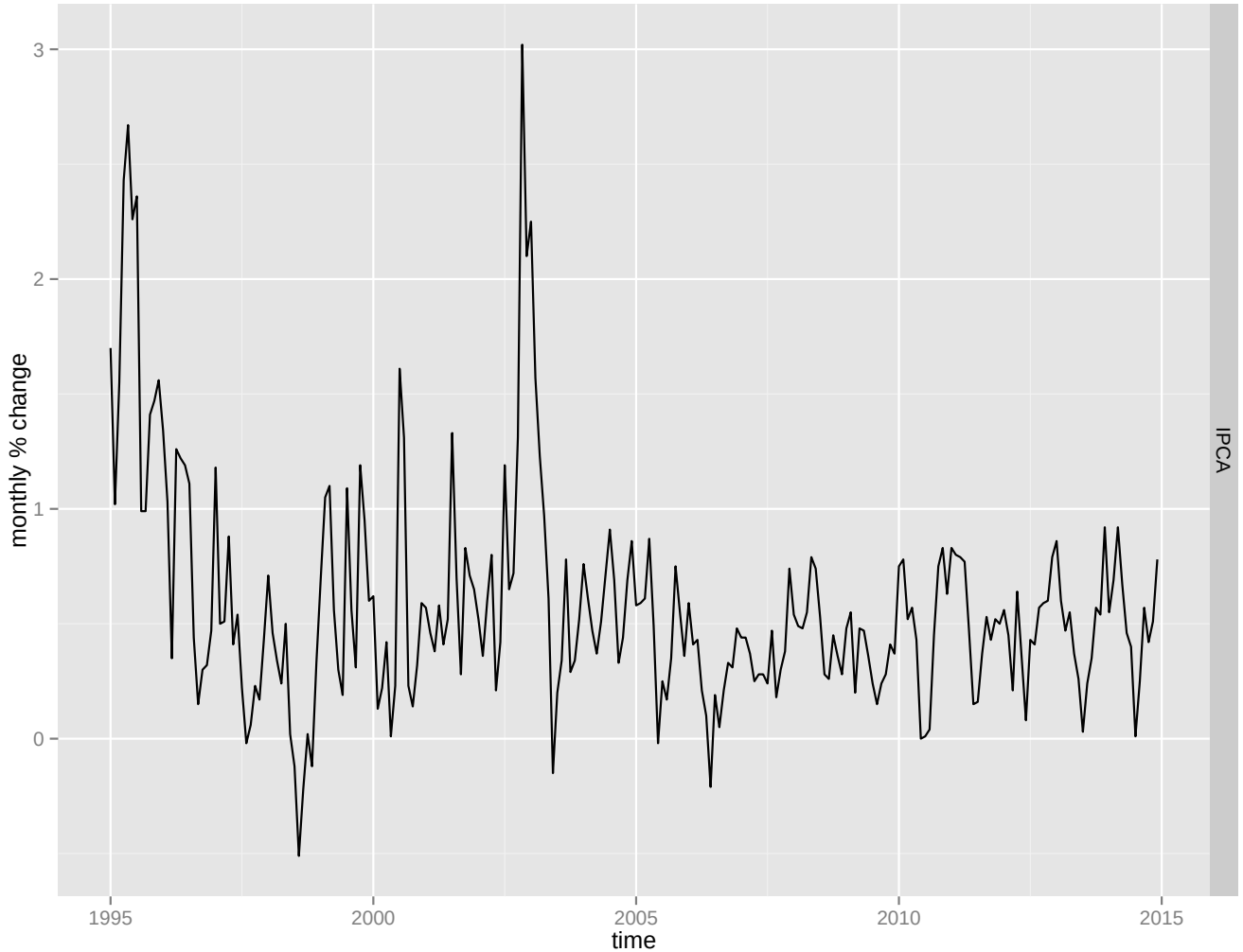
May 2015

Consumer Price Indexes: Dynamics dictated by "innovations" (shocks):

- regional (supply + demand)
- item-specific (supply)
- aggregate (demand)

Stochastic Index Number Theory: literature on Core Inflation and identification of orthogonal shocks. Here: explore the hierarchical structure of price movements along different dimensions.

Brazilian Broad Consumer Price Index (IPCA) covers 10 metropolitan regions and 3 cities. Low volatility regime sample: January 1995 to December 2014.



- Classes [8]
  - Items [52]
    - Sub-items [373]

In its current configuration, IPCA is divided into 8 classes (Foodstuff, Transportation, Household expenses, Health, Education, Clothing, Leisure and Communications), which are then divided into 52 Items), and finally in 373 sub-items. This will provide the hierarchical structure which will be explored here.

X_{ibt} items of the IPCA, i being the series subscript, b the block where the series belongs and t the particular moment in time.

G_b factor related to block b.

e_{Xibt} idiosyncratic variation of a particular series at a particular point in time.

Block factors G_{bt} are related to the aggregate factors F_t, which are in turn dynamically formulated as depending on their own past values, and on their own unexplained variations (interpreted as "innovations").

Using the notation where individual series and factors as grouped in vectors

$$X_{bt} = (X_{bt.1}, X_{bt.2}, \ldots, X_{bt.N_b})$$
$$G_{bt} = (G_{bt.1}, G_{bt.2}, \ldots, G_{bt.k_b})$$

the model can be concisely formulated including dynamics as

$$\begin{aligned}
X_{bt} &= \Lambda_{G.b0} G_{bt} + \ldots + \Lambda_{G.bs_{Gb}} G_{b,t-s_{Gb}} + e_{Xbt} \\
G_{bt} &= \Lambda_{F.b0} F_t + \ldots + \Lambda_{F.bs_F} F_{t-s_F} + e_{Gbt} \\
F_t &= \Psi_{F.1} F_{t-1} + \ldots + \Psi_{F.q_F} F_{t-q_F} + \epsilon_{Ft} \\
e_{Gbt} &= \Psi_{G.b1} e_{Gb,t-1} + \ldots + \Psi_{G.bq_{Gb}} e_{Gb,t-q_{Gb}} + \epsilon_{Gbt} \\
e_{Xbit} &= \Psi_{X.bi1} e_{Gb,t-1} + \ldots + \Psi_{X.biq_{Xb}} e_{Xbit,t-q_{Xb}} + \epsilon_{Xbit}
\end{aligned}$$

The error terms in all equations follow the usual distributional assumptions (normality, zero mean and constant variance).

The parameters in the matrices Λ and Ψ, along with the variances of the error terms in all equations are jointly estimated from data with the factors in all levels, in the context of a state-space model where a joint probability distribution is assumed for all these random quantities.
Details of inference using the Gibbs sampling algorithm for this model: Moench E. , Ng S. and Potter S. (2009).

The blocks are defined by IPCA items group in five categories, representing different economic dynamics for prices:

Foodstuff Food items consumed inside the house.

Non-regulated services Services provided by independent professionals, not subject to direct government intervention.

Regulated Prices Services either directly provided by the government, or whose prices are subject to direct intervention by the government.

Non and semi-durable consumer goods Goods not purchased on long-term credit.

Durable consumer goods Goods normally purchased on several installments or directly financed.

Outline I

Group	Item
Foodstuff	1101.Cereals and oilseeds
	1102.Flour, starches and pastas
	1103.Tubercles, roots and greenstuff
	1104.Sugar and derivatives
	1105.Vegetables
	1106.Fruits
	1107.Meats
	1108.Fish
	1109.Processed meats and fish
	1110.Poultry and eggs
	1111.Milk and derivatives

Outline II

	1112.Bakery 1113.Oils and fats 1114.Drinks 1115.Canned and preserved 1116.Salt and spices
Non-regulated services	1201.Meals (non-home) 2101.Rents and fees 2103.Repairs 3301.Maintenances 6201.Medical and dental services 6202.Hospital and lab services 7101.Personal services 7201.Leasure 8101.Regular courses

Outline III

	8104.General courses
Regulated Prices	2201.Fuels (home use) 2202.Electrical Energy (residential) 5101.Public transportation 5104.Fuels (automobiles) 6101.Pharmaceuticals 6203.Health plans 7202.Smokes 9101.Communications

Outline IV

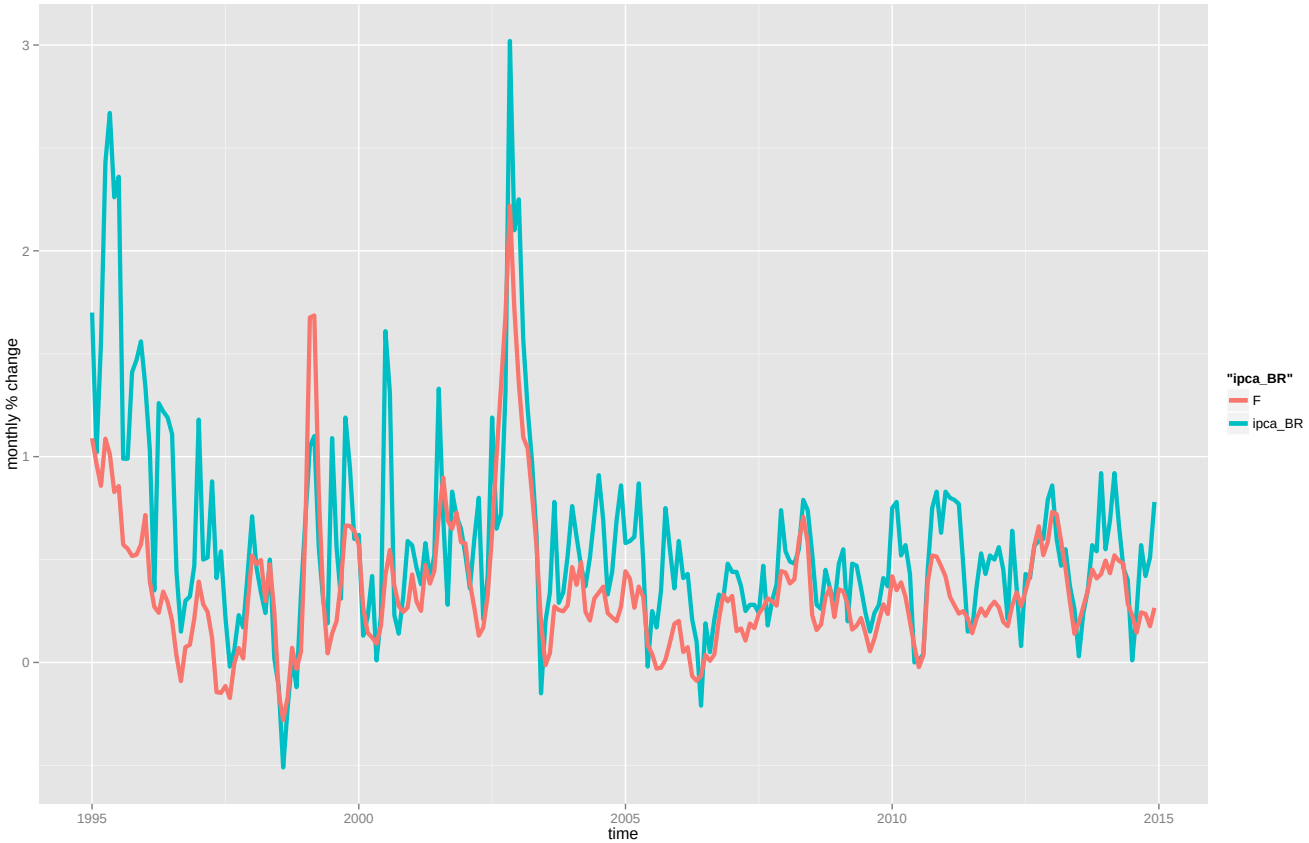
Non and semi-durable consumer goods	2104.Cleansing products 3102.Utensils and decorations 3103.Bed, table and bath cloths 4101.Menswear 4102.Womenswear 4103.Childrenswear 4201.Shoes and accessories 4301.Jewelery 4401.Cloths 6102.Optical products 6301.Personal hygiene 7203.Photography 8102.Books and magazines

Outline V

	8103.Stationery
Durable consumer goods	3101.Furniture 3201.Appliances 3202.TVs, stereo systems and computers 5102.Personal vehicles

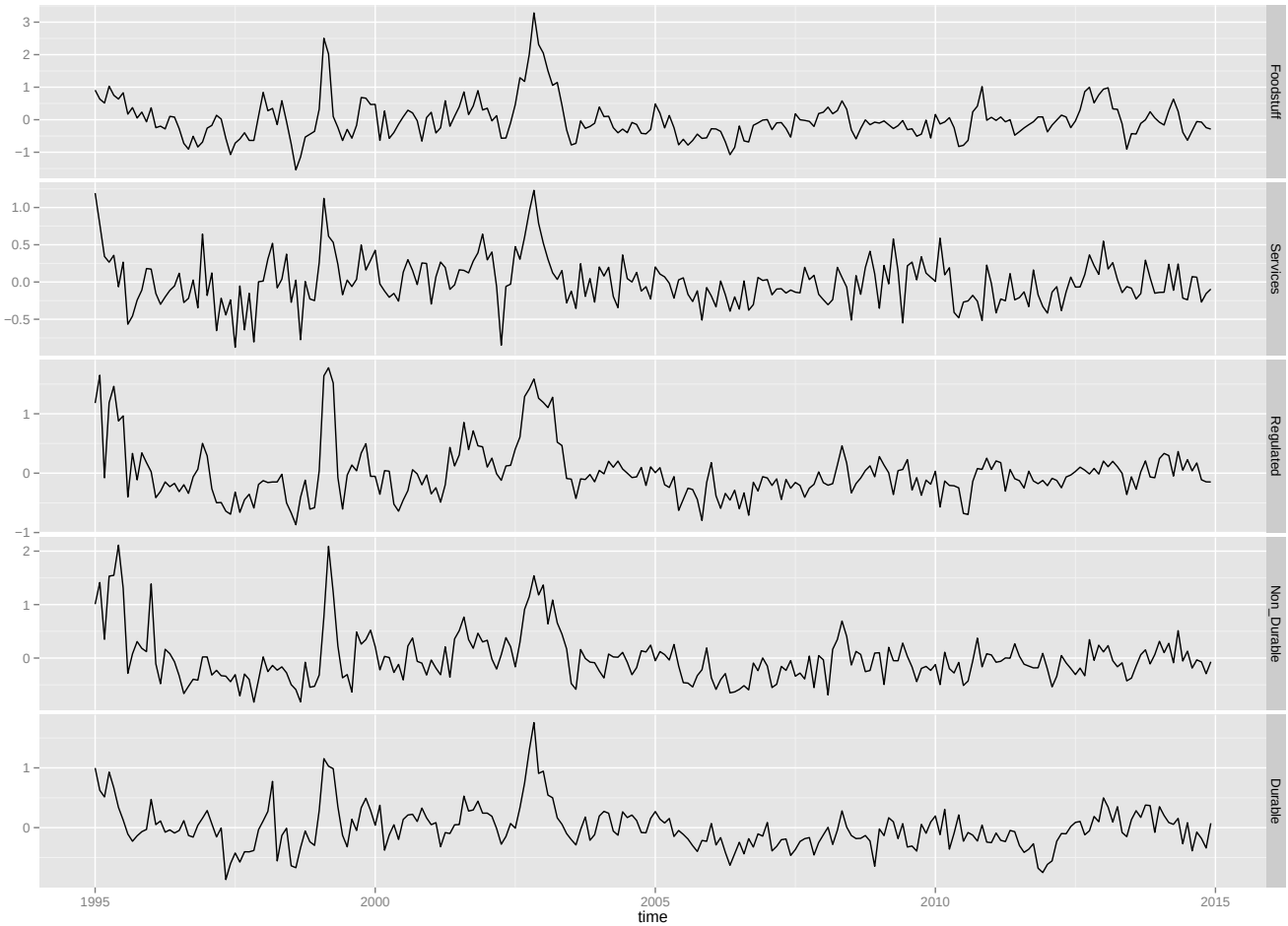
Outline VI

Factors are estimated along the hierarchy formed by the groups. Headline IPCA along with level-3 factor $\hat{F}_t$:



$\hat{F}_t$ can be interpreted as the commonality of behavior of price changes across groups, and as so provides an alternative measure of core inflation.
Striking feature: $\hat{F}_t$ is much smoother than the headline IPCA, and also leads its behavior on cyclical trends, both desirable features of this kind of measurement.

Estimated group factors $\hat{G}_1, \ldots, \hat{G}_5$:



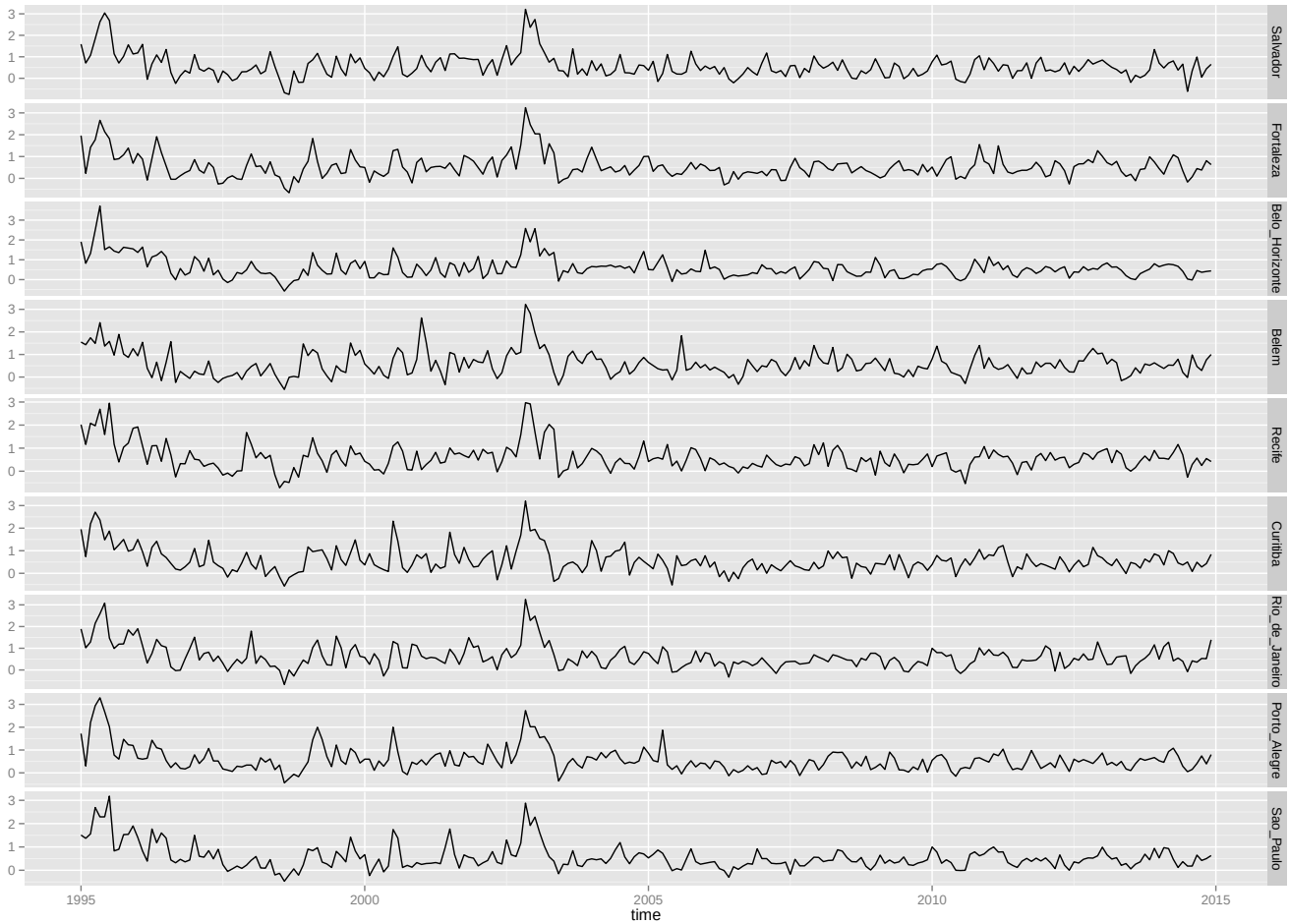
The structure of the IPCA panel also allows for further groupings of economic interest, such as the regional dimension. In terms of the above established notation, the individual series are now represented by Z_{bsit}, where the indices b, i and t have the same meanings as in the three levels model, but now augmented by s, the index for the sub-blocks.

Denoting H_{bst} as the vector of factors at the sub-block level, the model becomes

$$
\begin{aligned}
Z_{bsit} &= \lambda_{H.bsi}(L)H_{bst} + e_{Xbsit} \\
H_{bst} &= \Lambda_{G.bs}(L)G_{bt} + e_{Hbst} \\
G_{bt} &= \Lambda_{F.b}(L)F_t + e_{Gbt} \\
F_t &= \Psi_F(L)F_{t-1} + \epsilon_{Ft} \\
e_{Gbt} &= \Psi_{G.b1}e_{Gb,t-1} + \ldots + \Psi_{G.bq_{Gb}}e_{Gb,t-q_{Gb}} + \epsilon_{Gbt} \\
e_{Hbst} &= \Psi_{H.bs1}e_{Hbs,t-1} + \ldots + \Psi_{H.bsq_{Hbs}}e_{Hbs,t-q_{Hbs}} + \epsilon_{Hbst} \\
e_{Xbit} &= \Psi_{X.bi1}e_{Gb,t-1} + \ldots + \Psi_{X.biq_{Xb}}e_{Xbit,t-q_{Xb}} + \epsilon_{Xbit}
\end{aligned}
$$

Again, all error terms follow the usual distributional assumptions.

Estimated Regional Factors:



Each series in the four-level model can have its sample variance decomposed into four different sources of variances:

- $share_F$ Aggregate level.
- $share_G$ Block-level (regional).
- $share_G$ Sub-block level (economic groups).
- $share_Z$ Individual items.

Item	$share_F$	$share_G$	$share_H$	$share_Z$
Foodstuff	0.130	0.027	0.027	0.815
Services	0.012	0.003	0.194	0.792
Regulated	0.001	0.000	0.291	0.708
Consumer Non-durable	0.057	0.012	0.043	0.888
Consumer Durable	0.052	0.011	0.132	0.805

There appears to be a lot of useful information contained in the estimated factors of a multi-level model for inflation and relative price changes. The two results presented here can be extended on a number of dimensions:

Alternative groupings There are other types of aggregation of the index sub-items. For instance two groups which would overlap with the ones considered here are tradables vs. non-tradables, which factors could be correlated with macroeconomic fundamentals such as the exchange rate and salaries.

Dynamics The model is completely general in its possible specifications for lags between observed series and factor, between factors of different levels, and between the random variables behind the error terms. In parallel with results from time-series models that do not include latent variables, the results can be very sensitive to alternative specifications in the dynamics.

Number of Factors Like in one-level factor models, it is possible to estimate more than one factor for each level in the model adopted here. As already mentioned, the reason for adopting the hierarchical structure is that it provides a direct interpretation for each distinct factor. However, the information contained in the data may not be enough to exhaust the possibilities of meaningfully estimated latent variables. In this sense, more than one factor can be potentially estimated for each of the model's levels.

Data from different sources In principle, there is no restriction for including data on price changes calculated by different institutions in differents regions. This opens the possibility of defining new types of groupings and levels.

Thank you!