June 2006

Percentiles of an Inflation Index by Quantile Regression

Eric Blankmeyer
Department of Finance and Economics
McCoy College of Business Administration
Texas State University – San Marcos
San Marcos, TX 78666

Email eb01@txstate.edu

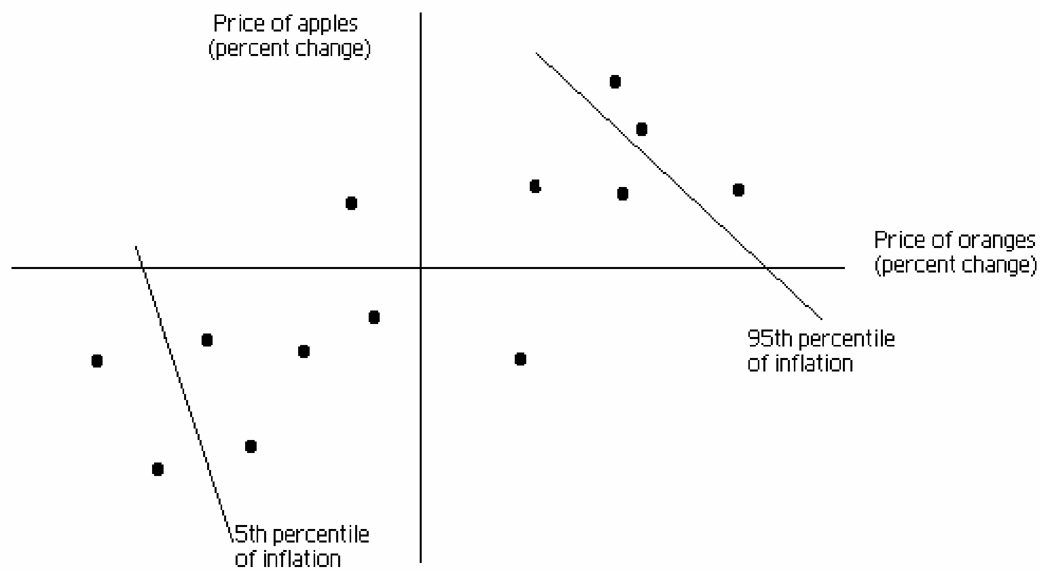
Abstract. This paper gives a methodology for estimating an inflation index using the quantile regression of Bassett and Koenker. The regression, which is orthogonal in the logarithmic price changes, is computed by linear programming for each percentile of inflation. The procedure is applied to monthly data on 25 raw materials.

Key words. Inflation, price index, quantile regression.

Percentiles of an Inflation Index by Quantile Regression

Let p_{ta} and p_{to} denote the percent changes in the prices of apples and oranges between period t-1 and period t, where t = 1, …, T. Several pairs (p_{ta} , p_{to}) are displayed in Figure 1.

Figure 1. Inflation percentiles from quantile regression



An index of fruit-price inflation requires weights w_a and w_o , non-negative numbers that sum to 1. Let us compute the weights as the coefficients in an orthogonal regression that minimizes the sum of the absolute values of the discrepancies between the observations and the line:

$$\text{Minimize } \Sigma_t \mid p_{ta}w_a + p_{to}w_o - \pi \mid \quad . \qquad (1)$$

The regression line has an intercept π and a slope $-w_o / w_a$; actually, we shall fit several such lines with different intercepts and possibly different slopes. In Figure 1, the left-hand line corresponds to the fifth percentile of inflation (say, π_{05}) while the right-hand line corresponds to the ninety-fifth percentile of inflation, π_{95} .

In general, the percentiles of an inflation index can be estimated by the method of quantile regression (Koenker and Hallock 2001). Given K commodities, consider the linear model

$$\Sigma_k \, p_{tk} w_k - \pi = u_t \quad , \qquad t = 1, \ldots, T. \tag{2}$$

Since the discrepancy u_t may be positive or negative, we define it as the difference between two non-negative numbers: $u_t = h_t - g_t$. The model becomes

$$\Sigma_k \, p_{tk} w_k - \pi = h_t - g_t \quad , \quad t = 1, \ldots, T \tag{3}$$

subject to h_t , g_t , $w_k \geq 0$ and $\Sigma_k w_k = 1$. Because $| u_t | = h_t + g_t$, we have a *linear programming problem* in which $\Sigma_t (h_t + g_t)$ is minimized subject to (3), the normalization of the weights, and the non-negativity constraints. In the optimal solution, π is an estimate of the fiftieth percentile, the median rate of inflation. More generally, we can minimize $\Sigma_t [\theta h_t + (1-\theta) g_t]$, in which case π estimates the θth percentile of inflation ($0 < \theta < 1$). Varying θ parametrically, we solve a sequence of linear-programming problems that yields the frequency distribution of the inflation index; with specialized algorithms, these computations can be performed very efficiently even for large values of T and K (Portnoy and Koenker 1997).

As part of the data preparation, a researcher may choose to scale p_{tk} ; for example, the observations on each commodity may be multiplied by its expenditure share in a representative household budget. Alternatively, additional linear constraints may be used to specify the total weight for each *group* of commodities (e. g. food, clothing, housing, and so on); these constraints could be expressed as equations or inequalities. (For the application discussed below, however, no such *a priori* scaling or grouping has been performed.) It is also worth mentioning that some of the weights w_k may turn out to be zero. This occurs when the price variations of two or more commodities are highly correlated: in the presence of near co-linearity, the quantile regression opts for parsimony by dropping a redundant commodity.

As an application, we compute an inflation index for a set of twenty-five raw materials including metals, fibers, cereals, vegetable oils, other foods, beverages, animal products and rubber. The prices are unit values from the International Monetary Fund (2004). We use monthly logarithmic changes of each commodity price from February 1975 to January 2004 (348 observations). Table 1 displays several percentiles of this commodity inflation index. While the median monthly price change is close to zero, there is considerable deflation at the fifth percentile and considerable inflation at the ninety-fifth percentile. The interquartile range, a robust measure of dispersion, is 2.77 percent per month [1.47 - (-1.30)]. Table 2 shows each commodity's weight by percentile. Cocoa, cotton, gold, rice, wheat and wool consistently have the largest weights. Barley, lead and sorghum are always omitted, while the weights for copper, copra and sugar are negligible. In conclusion, the orthogonal quantile regression appears to be a feasible and useful way to characterize the empirical frequency distribution of inflation in this set of commodity prices.

References

International Monetary Fund (2004). *International Financial Statistics Online Service*. http://imfstatistics.org/imf .

Koenker, Roger, and Kevin F. Hallock (2001). "Quantile Regression: An Introduction." *Journal of Economic Perspectives* 15 (4), 143-156.

Nyquist, Hans (2002). "Orthogonal L1-norm Estimation" in Yadolah Dodge (editor), *Statistical Data Analysis Based on the L1-Norm and Related Methods*. Basel: Birkhauser., 171-181.

Portnoy, Stephen, and Roger Koenker (1997). "The Gaussian hare and the Laplacian tortoise." *Statistical Science* 12, 279-300.

Table 1. Estimated percentiles of inflation

Percentile	Inflation (% / month)
Fifth	-2.94
Twenty-fifth	-1.30
Fiftieth	0.06
Seventy-fifth	1.47
Ninety-fifth	3.43

Table 2. Commodity weights by percentile of inflation

PERCENTILE	ALUMINUM	BARLEY	COCOA	COFFEE	COPPER
5TH	0.0170	0.0000	0.0948	0.0496	0.0000
25TH	0.0379	0.0000	0.0820	0.0128	0.0000
50TH	0.0279	0.0000	0.0630	0.0061	0.0000
75TH	0.0312	0.0000	0.0466	0.0078	0.0000
95TH	0.0238	0.0000	0.0681	0.0000	0.0373

PERCENTILE	COPRA	COTTON	GOLD	HIDES	LEAD
5TH	0.0000	0.0805	0.1195	0.0076	0.0000
25TH	0.0000	0.0591	0.1097	0.0782	0.0000
50TH	0.0000	0.0688	0.1136	0.0308	0.0000
75TH	0.0000	0.0753	0.0838	0.0480	0.0000
95TH	0.0005	0.1298	0.0253	0.0941	0.0000

PERCENTILE	LINDSEED	MAIZE	NICKEL	ORANGES	PALM
5TH	0.0182	0.0461	0.1089	0.0141	0.0000
25TH	0.0601	0.0700	0.0280	0.0172	0.0004
50TH	0.0133	0.0782	0.0000	0.0130	0.0243
75TH	0.0000	0.0282	0.0000	0.0218	0.0058
95TH	0.0000	0.0000	0.0390	0.0097	0.0297

PERCENTILE	RICE	RUBBER	SORGHUM	SOYBEANS	SUGAR
5TH	0.1745	0.0414	0.0000	0.0040	0.0000
25TH	0.1415	0.0152	0.0000	0.0032	0.0080
50TH	0.1344	0.0389	0.0000	0.0000	0.0042
75TH	0.1078	0.0351	0.0000	0.0000	0.0020
95TH	0.1139	0.0000	0.0000	0.0000	0.0000

PERCENTILE	TEA	TIN	WHEAT	WOOL	ZINC
5TH	0.0577	0.0167	0.0886	0.0609	0.0000
25TH	0.0367	0.0527	0.0518	0.1157	0.0201
50TH	0.0475	0.0586	0.0782	0.1592	0.0399
75TH	0.0586	0.0745	0.1411	0.1549	0.0775
95TH	0.0290	0.1385	0.1055	0.0963	0.0594