Journal of Economics Library
www.kspjournals.org
Volume 3 March 2016 Issue 1

The Path-Dependence Bias in Approximating Local Price Levels by CPIs

By Konstantin GLUSCHENKO †

Abstract. Lacking data on price levels across locations, economists are forced to proxy them. One method is to extrapolate the price levels known for locations in some point in time to another point by multiplying the initial price levels by the local CPIs. With the use of simulation experiments, this paper demonstrates that such a method is inadequate, since the path dependence of CPI alone produces considerable biases distorting cross-location comparisons of price levels.
Keywords. Spatial price index, Inflation; Divisia index, Nonhomothetic preferences, Demand system.
JEL. C15, C43**,** E31, R19.

1. Introduction

To correctly compare monetary indicators across locations (countries, national regions, cities, etc.) in real terms, economists need data on local price levels. Commonly, relative price levels are dealt with, taking some location as the base, which gives spatial price indexes (SPI). Lacking such data, economists resort to local consumer price indexes (CPI) to estimate SPIs. That is, provided that local price levels are known at some base point in time, an economist extrapolates them to a given point with the use of changes in the price levels, i.e. CPIs.

To name a few, Chen & Devereux (2003) exploit this procedure to construct price levels for US cities; Solanko (2008) estimates real incomes across Russian regions through extrapolated regional price levels; and Faber & Stockman (2009) use EU's Harmonized Index of Consumer Prices to asses price levels in European countries. Estimation of purchasing power parities (PPPs) for non-survey years also bases on the use of national CPIs as extrapolation factors (Eurostat & OECD, 2012, p. 132). Such an approach seems doubtful for two main reasons. First, there is a conceptual inconsistency between spatial and temporal price indexes. Second, the CPI is known to suffer from a number of biases which are hardly uniform across locations. Therefore, one may reasonably expect the CPI-extrapolated SPIs to be biased, thus distorting spatial comparisons.

Even elimination of these two concerns does not save the situation. The bias is unavoidable if for no other reason than the path dependence of CPI. This paper uses simulated data to understand the extent of distortions caused by this reason within a simple two-good two-location framework. Consider a time span $t = 0,\ldots,T$. At starting point $t = 0$, prices $p_{kr}(t)$ are set equal both across goods ($k = X$,

† Institute of Economics and Industrial Engineering, Siberian Branch of the Russian Academy of Sciences, Russia; and Novosibirsk State University, Russia.
☎. +7 383 330 2548
✉. glu@nsu.ru

Y) and locations (r = 1, 2). They randomly change but eventually return to equal values at final point $t = T$ as displayed in Figure 1. Thus, the 'actual' SPI, $P_{12}(t) = (W_{X1}(t)p_{X1}(t) + (1 - W_{X1}(t))\, p_{Y1}(t))/(W_{X2}(t)p_{X2}(t) + (1 - W_{X2}(t))\, p_{Y2}(t))$, W standing for weights, equals 1 at time points 0 and T under any definition of the weights.

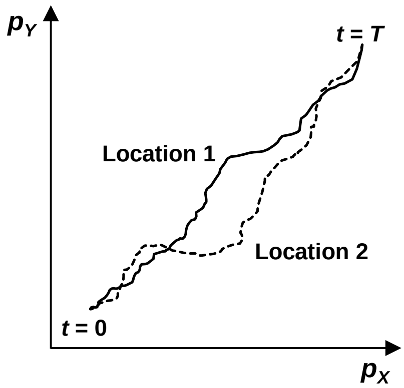

Figure 1. *Price paths*

Representative consumers are identical across locations, having the same preferences and nominal incomes. This allows to get rid of the first problem, the conceptual inconsistency between spatial and temporal price indexes, as we can deem a consumer from any location to be an 'over-location' representative who can equally well confront prices at both locations. To compute the CPI-extrapolated price level, the Divisia price index (see, e.g. Hulten, 2008) is applied. This allows to get rid of biases in CPIs, the formula bias among them. With continuously changing weights, the Divisia index is the most exact CPI; all other formulas of chained CPI are approximations of it.

Under these conditions, the sole source of bias in the extrapolated SPI is the path dependence of CPI. That is, despite changes in prices themselves over time T are equal across locations, local CPIs prove to be unequal due to different price paths. It is worth noting that the CPIs themselves cannot be deemed biased, as the path dependence is an inherent property of measuring price level changes by chained indexes. (The only case of path invariance is that of homothetic preferences – Samuelson & Swamy, 1974; however, it is unrealistic, implying unity income elasticity of demand for all commodities.) Comparing the 'actual' and CPI-extrapolated SPI at T, a bias in the latter is estimated. Generating a great number of random price paths yields a distribution of the bias. Results obtained suggest that the path dependence of CPI alone produces considerable biases distorting cross-location ratios of price levels.

2. Design of simulations

Simulating price dynamics. Prices are continuous time functions. Let π_{rkt} be a change in price for good k in location r (percentage price shock) over a unit time interval $[t - 1, t]$, i.e. $p_{kr}(t) = (1 + \pi_{krt})\cdot p_{kr}(t - 1)$; $p_{kr}(0) = p_0$. Within intervals $[t - 1, t]$, the changes are linear: $p_{kr}(t - 1 + \tau) = (1 + \pi_{krt}\tau)\cdot p_{kr}(t - 1)$, $\tau \in [0, 1]$. Price shocks are random and independent across locations and goods, but they depend on their own past values through an autoregressive process AR(1): $\pi'_{krt} = \rho\pi'_{kr,t-1} + \varepsilon_{krt}$, $\pi'_{kr0} = 0$, $\pi'_{kr} > -1$ (otherwise it is reestimated; no one such event occurred during the simulations), where π'_{krt} is a 'raw' (nonnormalized) value of price shock, ρ is an autoregressive coefficient ($0 \le \rho \le 1$), and ε_{krt} is i.i.d. $N^*(0, \sigma_-^2, \sigma_+^2)$. Distribution $N^*(0, \sigma_-^2, \sigma_+^2)$ is a right skewed one in order to make price-cutting less likely than rise in prices: $\varepsilon_{krt} < 0$ are drawn from $N^{-}(0, \sigma_-^2)$ with probability $\sigma_-/(\sigma_- + \sigma_+)$, and $\varepsilon_{krt} \geq 0$ are drawn from $N^{+}(0, \sigma_+^2)$ with probability $\sigma_+/(\sigma_- + \sigma_+)$; $\sigma_- < \sigma_+$. Normalization $\pi_{krt} = (1+\pi'_{krt})(1+\bar{\pi})\prod_{z=1}^{T}(1+\pi'_{krz})^{-1/T} - 1$ ensures the geometric average (over the whole time span) of rises in prices, $1 + \pi_{rkt}$, to be uniform for all goods and locations and equal to $1+\bar{\pi}$ (where $\bar{\pi}$ is a predetermined value). Hence, the cumulative rise in all prices becomes the same at $t = T$, equaling $(1+\bar{\pi})^T$.

Incomes. Nominal incomes, $m_r(t)$, are the same in both locations, $m_1(t) = m_2(t) = m(t)$. In contrast to prices, incomes change discretely, remaining constant within unit time intervals $(t–1, t]$, i.e. $m(t-1+\tau) = m_t$, $\tau \in (0, 1]$. They steadily change with a constant rate: $m_t = m_0((1+\bar{\pi})^T n)^{(t-1)/(T-1)}$, so that the real income at the final point $t = T$ equals $n{\cdot}m_0$. Thus, depending on whether n is more or less than unity, real incomes may either rise or fall with time.

Modeling consumption. One or another of three demand systems model consumer behavior. Suppressing the location and time subscripts to economize notations and denoting the quantity of good k by q_k, these demand systems look like

$$\begin{cases} q_X = q_{X0} + \alpha(m - q_{X0}p_X)/p_X \\ q_Y = (1-\alpha)(m - q_{X0}p_X)/p_Y \end{cases}, \quad (1)$$

$$\begin{cases} q_X = q_{X0}(1+\log(m/p_X q_{X0})) \\ q_Y = (m - q_{X0}(1+\log(m/p_X q_{X0}))p_X)/p_Y \end{cases}, \quad (2)$$

$$\begin{cases} q_X = \beta q_{X0}(1-\exp(-\delta m/p_X q_{X0})) \\ q_Y = (m - \beta q_{X0}(1-\exp(-\delta m/p_X q_{X0}))p_X)/p_Y \end{cases}. \quad (3)$$

All three systems assume a nonzero minimum consumption of X, q_{X0} (say, X designates foods and Y non-foods), and $m \geq p_X q_{X0}$ to hold. In the above formulas, α and β are parameters; $0 < \alpha < 1$; $\beta > 1$; $\delta = \log(\beta/(\beta-1))$.

Different nonhomothetic preferences generate the above demand systems. The Stone-Geary preferences $U(q_X, q_Y) = (q_X - q_{X0})^{\alpha} q_Y^{1-\alpha}$ yield demand system (1). Preferences of the form $U(q_X, q_Y) = q_Y \exp\left(\int^{q_X} \frac{dq}{q_{X0}\exp(q/q_{X0}-1) - q}\right)$ imply that the income elasticity of demand for X asymptotically tends to zero with increasing quantity of X: $\varepsilon_{mX} = q_{X0}/q_X$. This gives demand system (2). At last, in preferences $U(q_X, q_Y) = q_Y \exp\left(-\int^{q_X} \frac{dq}{q_{X0}\log(1-q/\beta q_{X0})/\delta + q}\right)$, consumption of X is assumed to have a saturation level βq_{X0}, approached as $m/p_X \to \infty$. We obtain herefrom demand system (3).

Computing CPI. In the standard manner, expenditure equals income, $q_{Xr}(t)p_{Xr}(t) + q_{Yr}(t)p_{Yr}(t) = m(t)$. A CPI over $[0, T]$ for location r is computed as the Divisia price index $I_r(0,T) = \exp(\int_0^T \frac{q_{Xr}(t)\cdot dp_{Xr}(t)/dt + q_{Yr}(t)\cdot dp_{Yr}(t)/dt}{m(t)}dt)$. Given that prices are piecewise-linear functions of time and expenditures are piecewise constant, it takes the form

$$I_r(0,T) = \exp(\sum_{t=1}^{T}(\frac{p_{Xr,t-1}\pi_{Xrt}}{m_t}\int_0^1 q_{Xr}(p_{Xr}(t-1+\tau), p_{Yr}(t-1+\tau))d\tau + \frac{p_{Yr,t-1}\pi_{Yrt}}{m_t}\int_0^1 q_{Yr}(p_{Xr}(t-1+\tau), p_{Yr}(t-1+\tau))d\tau, \quad (4)$$

where $p_{kr,t-1} = p_0\cdot(1+\pi_{kr1})\cdot\ldots\cdot(1+\pi_{kr,t-1})$; recall that all initial prices are equal. To compute (4), numerical integration is implemented.

For comparison, a CPI similar to that employed by most national statistical agencies, the chained Laspeyres-type index, is also computed. (In fact, this is the Lowe index rather than the original Laspeyres index – see ILO et al., 2004, pp. 2–3.) For two goods, a one-period index looks like $I_r(t-1,t) = w_{Xr\theta}\frac{p_{Xr}(t)}{p_{Xr}(t-1)} + w_{Yr\theta}\frac{p_{Yr}(t)}{p_{Yr}(t-1)}$, where θ is a weight reference period; the chained CPI over $[0, T]$ is $I_r(0, T) = I_r(0, 1)\cdot\ldots\cdot I_r(T-1, T)$. Weights $w_{kr\theta}$ are updated 'yearly,' based on the expenditure pattern over the previous 'year.' Then θ relates to that 'year,' being calculated as $\lfloor(t-1)/12\rfloor$, where $\lfloor x\rfloor$ stands for the integer part ('floor') of x. For $\theta \geq 1$, $w_{Xr\theta} = \sum_{z=1}^{12} q_{Xr}(12(\theta-1)+z)p_{Xr}(12(\theta-1)+z) / \sum_{z=1}^{12} m(12(\theta-1)+z)$; for $\theta = 0$, $w_{Xr0} = q_X(p_0)p_0/m_0$; $w_{Yr\theta} = 1 - w_{Xr\theta}$.

Estimating SPIs and the bias. Let $P_r(t)$ be an actual price level. Then the ratio $P_{12}(t) = P_1(t)/P_2(t)$ gives the actual SPI. The CPI-extrapolated price level looks like $P'_r(t) = P_r(0)\cdot I_r(0, t)$. Call $P'_{12}(t) = P'_1(t)/P'_2(t) = P_{12}(0)\cdot I_1(0, t)/I_2(0, t)$ the indirect SPI. Since $P_{12}(0) = 1$ by construction, $P'_{12}(T) = I_1(0, T)/I_2(0, T)$. Its deviation from the actual SPI, $(P'_{12}(T) - P_{12}(T))/P_{12}(T)$, estimates the bias in indirect SPI relative to the actual one. As – also by construction – $P_{12}(T) = 1$, the bias is equal to $P'_{12}(T) - 1$.

3. Results

The results reported below are obtained for $T = 120$ (10 'years' × 12 'months'). The average 'monthly' price shock, $\bar{\pi}$, equals 1.35%, yielding 'annual' rise in prices of 17.5% and a fivefold rise in prices over the whole time span (such a figure is not extraordinary, e.g., inflation in Turkey over 2000–2009 increased the overall price level by a factor of 5.62); $\sigma_- = 0.1\bar{\pi}$ and $\sigma_+ = 1.4\bar{\pi}$; $\rho = 0.5$. Figure 2 depicts a kernel estimate of the distribution of simulated price shocks π_{krt}.

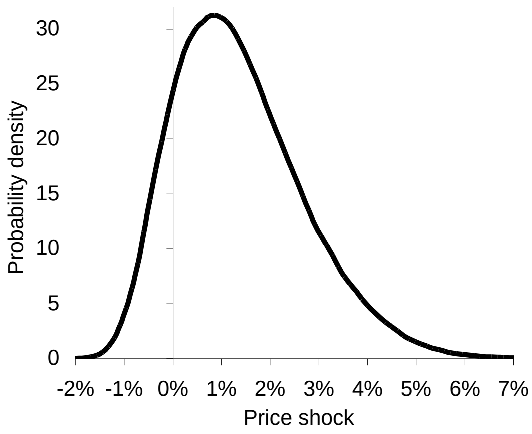

Figure 2. *Distribution of simulated price shocks*

The number of replications is 10,000 in each experiment. Parameters of the demand systems are: $\alpha = 0.1$, $\beta = 1.5$, and $q_{X0} = 0.9$. Starting prices are $p_{rk0} = 1$; final prices are $p_{krT} = p_0(1.0135)^{120} \approx 5p_0$. Nominal incomes are set in two ways that provide rising and falling real income. First, incomes rise from $m_0 = 1$ with 'monthly' rate about 2.14% to $m_T \approx 12.5$, thus, real incomes at T are 2.5 times higher than at $t = 0$. Second, $m_0 = 2$, 'monthly' rate is about 0.77%, and $m_T \approx 5$, final real incomes becoming half as much as the initial ones.

Figure 3 summarizes results obtained, reporting kernel estimates of the distribution of biases in indirect SPIs. Each panel of the figure corresponds to one of three demand systems; it demonstrates results for the cases of rising and falling real incomes and for two methods of extrapolating the indirect SPI, namely with the use of the Divisia and Laspeyres CPI. Table 1 reports summary statistics of the distributions.

Table 1. *Summary statistics of distributions of biases in indirect SPIs (%)*

Demand system	Index used for indirect SPI	Real income	Mean	Minimum	Maximum	Standard deviation
(1)	Divisia	rising	0.4	–24.3	30.2	6.8
	Laspeyres		0.5	–23.3	30.7	7.1
	Divisia	falling	0.2	–26.0	23.7	5.8
	Laspeyres		0.1	–28.9	25.8	5.6
(2)	Divisia	rising	0.1	–13.1	18.5	4.1
	Laspeyres		0.1	–13.2	18.8	3.8
	Divisia	falling	0.0	–10.2	12.0	2.8
	Laspeyres		0.0	–11.6	10.6	2.9
(3)	Divisia	rising	0.2	–24.0	26.0	6.3
	Laspeyres		0.2	–23.0	27.2	6.2
	Divisia	falling	0.1	–16.1	20.4	4.8
	Laspeyres		0.1	–15.6	21.8	4.8

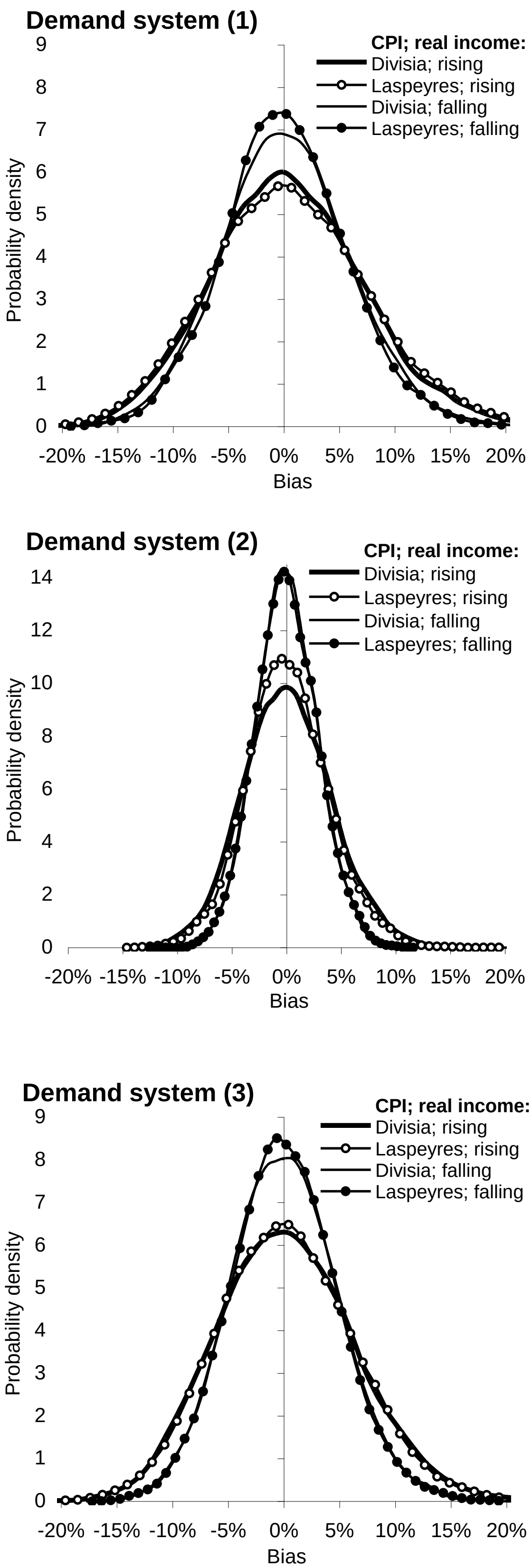

Figure 3. *Distributions of biases in indirect SPI*

These results indicate that the path dependence of CPI alone can sufficiently bias CPI-extrapolated SPIs. The most impressive are ranges of biases in Table 1, suggesting that the indirect SPI might be biased by up to 30% in either direction as compared to the direct SPI. Dispersion of biases (measured by standard deviations) is large as well, varying across different kinds of experiments from 3% to 7%.

The distributions of biases prove to be nearly symmetric around zero. Hence, estimates of SPI by extrapolation with the use of CPI can be either understated or overstated with approximately equal probability. The shapes of the distributions are roughly similar across demand systems. This provides hope that the pattern is qualitatively similar to what is actually occurring in the real world, whatever a real demand system may be.

Although we know the Laspeyres index is biased compared to the Divisia index due to the substitution effect, the distributions of the biases in indirect SPIs obtained with the use of Divisia and Laspeyres indexes are surprisingly close to each other. A possible explanation is that the substitution biases differ little between locations 1 and 2. Therefore, they almost cancel out in the indirect SPI which is the ratio of location CPIs.

The experiments not reported here may be summarized as follows. The higher and the more volatile inflation, the greater biases of indirect SPI (i.e. their standard deviation). This is valid for increases in both the average price shock, $\bar{\pi}$, and cumulative inflation with widening the time horizon T at a fixed $\bar{\pi}$. Volatility of inflation rises with increasing σ_{-}, σ_{+}, and/or ρ. The effect of random changes in nominal incomes instead of deterministic ones is similar to that of increasing volatility of inflation, enlarging – *ceteris paribus* – biases in indirect SPI.

4. Conclusions

The approach of approximating local price levels with the use of local CPIs is fairly common. The main conclusion of the simulation experiments is that such a procedure is biased even within a simple two-good two-location framework assuming identical preferences and nominal incomes of representative consumers in both locations. In reality, the pattern is much more complex. Actual CPIs cover a few hundreds of commodities with their own price paths; locations differ in income dynamics and preferences, etc. Therefore it may be expected that actual biases are much higher than those in our numerical experiments, being due not only to the path dependence. For instance, Gluschenko (2006), p. 22, finds indirect SPI to be biased across regions of Russia in the range of –8.1% to 10% over only 12 months (inflation equaling 10.1% over these 12 months).

Cross-country tests of the PPP also rely on approximating country price levels by national CPIs that can differ even in the commodity coverage. This seems to be one more clue to the 'PPP puzzle' posed by Rogoff (1966). A failure of time-series testing PPP may be an artifact caused by biases in relative CPIs involved, and not the result of price behavior.

Acknowledgements

This work was done while I was visiting the Bank of Finland Institute for Economies in Transition (BOFIT). Their hospitality is gratefully acknowledged.

References

Chen, L.-L., & Devereux, J. (2003). What can US city price data tell us about purchasing power parity? *Journal of International Money and Finance* 22(2), 213–222. doi. 10.1016/S0261-5606(02)00102-X

Eurostat & OECD (2012). *Eurostat-OECD Methodological Manual on Purchasing Power Parities*. Luxembourg: Publications Office of the European Union.

Faber, R. P., & Stockman, A. C. J. (2009). A short history of price level convergence in Europe. *Journal of Money, Credit and Banking* 41(2–3), 461–477. doi. 10.1111/j.1538-4616.2009.00215.x

Gluschenko, K. (2006). Biases in cross-space comparisons through cross-time price indexes: the case of Russia. BOFIT Discussion Paper No. 9.

Hulten, C. R. (2008). Divisia index. In: Durlauf, S.N., & Blume, L.E. (Eds.), *The New Palgrave Dictionary of Economics*. 2nd ed. Palgrave Macmillan.

ILO, IMF, OECD, UNECE, Eurostat, & World Bank (2004). *Consumer Price Index manual: Theory and practice*. Geneva: International Labour Office.

Rogoff, K. (1996). The purchasing power parity puzzle. *Journal of Economic Literature* 34 (2), 647–668.

Samuelson, P.A., & Swamy, S. (1974). Invariant economic index numbers and canonical duality: survey and synthesis. *American Economic Review* 64 (4), 566–593.

Solanko, L. (2008). Unequal fortunes: a note on income convergence across Russian regions. *Post-Communist Economies* 20(3), 287–301. doi: 10.1080/14631370802281399