Lumpy Price Adjustments: Supplements

Emmanuel Dhyne* **Catherine Fuss†** **Hashem Pesaran‡**
Patrick Sevestre§

March 9, 2009

Abstract

Supplemental Materials to "Lumpy Price Adjustments: A Microeconometric Analysis". Not to be published

*Banque Nationale de Belgique and Université de Mons-Hainaut.
†Banque Nationale de Belgique and Université Libre de Bruxelles.
‡Cambridge University, Faculty of Economics and CIMF.
§Paris School of Economics, Université Paris 1 - Panthéon Sorbonne and Banque de France.

A: Mathematical Proofs

Proof of the first part of Lemma 3.1.

$$E\left[yI(y+a)\right] = \sigma\phi\left(\frac{a+\mu}{\sigma}\right) + \mu\Phi\left(\frac{a+\mu}{\sigma}\right).$$

■

$$\begin{aligned} E\left[yI\left(y+a\right)\right] &= \int_{-a}^{+\infty} y\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy \\ &= \int_{-a}^{+\infty} \frac{y-\mu}{\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy + \int_{-a}^{+\infty}\frac{\mu}{\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy \end{aligned}$$

Letting $z = (y-\mu)/\sigma$, the above expression becomes

$$\begin{aligned} E\left[yI\left(y+a\right)\right] &= \sigma\int_{-\frac{a+\mu}{\sigma}}^{+\infty} z\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz + \mu\int_{-\frac{a+\mu}{\sigma}}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz \\ &= \sigma\left[-\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\right]_{-\frac{a+\mu}{\sigma}}^{+\infty} + \mu\int_{-\infty}^{\frac{a+\mu}{b}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz \\ &= \sigma\phi\left(\frac{a+\mu}{\sigma}\right) + \mu\Phi\left(\frac{a+\mu}{\sigma}\right) \end{aligned}$$

Proof of the second part of Lemma 3.1.

$$E\left[\phi\left(\frac{y+a}{b}\right)\right] = \frac{b}{\sqrt{b^2+\sigma^2}}\phi\left(\frac{a+\mu}{\sqrt{b^2+\sigma^2}}\right)$$

■

$$\begin{aligned} E\left[\phi\left(\frac{y+a}{b}\right)\right] &= \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y+a}{b}\right)^2}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy \\ &= \frac{1}{\sigma^2\pi}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{\left(\sigma^2+b^2\right)y^2+\left(2a\sigma^2-2b^2\mu\right)y+a^2\sigma^2+b^2\mu^2}{b^2\sigma^2}\right)}dy \\ &= \frac{1}{\sigma^2\pi}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{\left(\sqrt{\sigma^2+b^2}y+A\right)^2-A^2+a^2\sigma^2+b^2\mu^2}{b^2\sigma^2}\right)}dy \end{aligned}$$

where $A = (a\sigma^2 - \mu b^2) / \sqrt{b^2 + \sigma^2}$. Let $B = \frac{1}{2}\left(\frac{A^2 - a^2\sigma^2 - b^2\mu^2}{b^2\sigma^2}\right) = -\frac{1}{2}\frac{(a+\mu)^2}{b^2+\sigma^2}$,

$$E\left[\phi\left(\frac{y+a}{b}\right)\right] = \frac{1}{\sigma^2\pi}e^B \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{\left(\sqrt{\sigma^2+b^2}y+A\right)^2}{b^2\sigma^2}\right)} dy$$

$$= \frac{1}{\sigma^2\pi}e^B \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\frac{\sigma^2+b^2}{b^2\sigma^2}\left(y+\frac{a\sigma^2-\mu b^2}{b^2+\sigma^2}\right)^2} dy$$

Setting $\omega = b\sigma/\sqrt{b^2+\sigma^2}$ and $\widetilde{\mu} = -\left(a\sigma^2 - \mu b^2\right) / \left(b^2+\sigma^2\right)$, we now have

$$E\left[\phi\left(\frac{y+a}{b}\right)\right] = \frac{1}{\sigma^2\pi}e^B \int_{-\infty}^{+\infty} e^{-\frac{1}{2\omega^2}(y-\widetilde{\mu})^2} dy$$

$$= \frac{1}{\sigma^2\pi}e^B\omega\sqrt{2\pi} = \frac{b}{\sqrt{b^2+\sigma^2}}\frac{1}{\sqrt{2\pi}}e^B$$

$$= \frac{b}{\sqrt{b^2+\sigma^2}}\phi\left(\frac{a+\mu}{\sqrt{b^2+\sigma^2}}\right)$$

Proof of the third part of Lemma 3.1.

$$E\left(\Phi\left(\frac{y+a}{b}\right)\right) = \Phi\left(\frac{a+\mu}{\sqrt{b^2+\sigma^2}}\right)$$

■

$$E\left[\Phi\left(\frac{y+a}{b}\right)\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{\frac{y+a}{b}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}w}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dwdy$$

Stating that $\frac{z+y+a}{b} = w$, the expression above becomes

$$E\left[\Phi\left(\frac{y+a}{b}\right)\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{0} \frac{1}{b\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{z+y+a}{b}\right)^2}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dzdy$$

$$= \int_{-\infty}^{0}\frac{1}{b}\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{z+y+a}{b}\right)^2}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dydz$$

$$= \int_{-\infty}^{0}\frac{1}{b}E\left[\phi\left(\frac{y+a+z}{b}\right)\right] dz$$

Using the second part of Lemma 1,

$$\begin{aligned} E\left[\Phi\left(\frac{y+a}{b}\right)\right] &= \int_{-\infty}^{0} \frac{1}{b}\frac{b}{\sqrt{b^2+\sigma^2}}\phi\left(\frac{z+a+\mu}{\sqrt{b^2+\sigma^2}}\right)dz \\ &= \frac{1}{\sqrt{b^2+\sigma^2}}\int_{-\infty}^{0}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{z+a+\mu}{\sqrt{b^2+\sigma^2}}\right)^2}dz \end{aligned}$$

Setting $(z+a+\mu)/\sqrt{b^2+\sigma^2}=\widetilde{z}$,

$$\begin{aligned} E\left[\Phi\left(\frac{y+a}{b}\right)\right] &= \frac{1}{\sqrt{b^2+\sigma^2}}\int_{-\infty}^{\frac{a+\mu}{\sqrt{b^2+\sigma^2}}}\frac{\sqrt{b^2+\sigma^2}}{\sqrt{2\pi}}e^{-\frac{1}{2}\widetilde{z}^2}d\widetilde{z} \\ &= \Phi\left(\frac{a+\mu}{\sqrt{b^2+\sigma^2}}\right) \end{aligned}$$

Proof of the uniqueness of $\tilde{f}_t$ (the non-linear cross section average estimator of f_t). Let

$$z_{it}(f_t)=\frac{d_{it}}{\sqrt{\sigma_s^2+\sigma_\xi^2}},$$

and

$$\begin{aligned} \widetilde{\Delta p}_{it} &= \frac{\Delta p_{it}}{\sqrt{\sigma_s^2+\sigma_\xi^2}},\ \tilde{\eta}_{it}=\frac{\eta_{it}}{\sqrt{\sigma_s^2+\sigma_\xi^2}}, \\ \widetilde{s} &= \frac{s}{\sqrt{\sigma_s^2+\sigma_\xi^2}}\geq 0,\ \delta^2=\frac{\sigma_\xi^2}{\sigma_s^2+\sigma_\xi^2}<1, \end{aligned}$$

and note that we have

$$\begin{aligned} \widetilde{\Delta p}_{it} &= z_{it}(f_t)+z_{it}(f_t)\left[\Phi\left(z_{it}(f_t)-\widetilde{s}\right)-\Phi\left(z_{it}(f_t)+\widetilde{s}\right)\right] && (1) \\ &\quad +\delta^2\left[\phi\left(z_{it}(f_t)-\widetilde{s}\right)-\phi\left(z_{it}(f_t)+\widetilde{s}\right)\right]+\tilde{\eta}_{it}. && (2) \end{aligned}$$

The cross-sectional average estimate of f_t is now given by the solution of the non-linear equation

$$\begin{aligned} \Psi(\tilde{f}_t) &= \sum_{i=1}^{N} w_{it}\{z_{it}(\tilde{f}_t)+z_{it}(\tilde{f}_t)\left[\Phi\left(z_{it}(\tilde{f}_t)-\widetilde{s}\right)-\Phi\left(z_{it}(\tilde{f}_t)+\widetilde{s}\right)\right] && (3) \\ &\quad +\delta^2\left[\phi\left(z_{it}(\tilde{f}_t)-\widetilde{s}\right)-\phi\left(z_{it}(\tilde{f}_t)+\widetilde{s}\right)\right]\}-a_{Nt} && (4) \\ &= 0, && (5) \end{aligned}$$

where $a_{Nt}=\sum_{i=1}^{N} w_{it}\widetilde{\Delta p}_{it}$.

First it is clear that $\Psi(\tilde{f}_t)$ is a continuous and differentiable function of f_t, and it is now easily seen that

$$\lim_{f_t \to +\infty} \Psi(\tilde{f}_t) \to +\infty \quad \text{and} \quad \lim_{f_t \to -\infty} \Psi(\tilde{f}_t) \to -\infty.$$

Also the first derivative of $\Psi(f_t)$ is given by[1]

$$\Psi'(\tilde{f}_t) = \frac{1}{\sqrt{\sigma_s^2 + \sigma_\xi^2}} \sum_{i=1}^{N} w_{it} q_{it},$$

where

$$q_{it} = 1 + \Phi\left(z_{it}(\tilde{f}_t) - \widetilde{s}\right) - \Phi\left(z_{it}(\tilde{f}_t) + \widetilde{s}\right) + (1-\delta^2) h(z_{it}(\tilde{f}_t)),$$

and

$$h(z_{it}(\tilde{f}_t)) = z_{it}(\tilde{f}_t)\left[\phi\left(z_{it}(\tilde{f}_t) - \widetilde{s}\right) - \phi\left(z_{it}(\tilde{f}_t) + \widetilde{s}\right)\right].$$

But since $1 - \Phi\left(z_{it}(\tilde{f}_t) + \widetilde{s}\right) = \Phi\left(-z_{it}(\tilde{f}_t) - \widetilde{s}\right)$, then

$$1 + \Phi\left(z_{it}(\tilde{f}_t) - \widetilde{s}\right) - \Phi\left(z_{it}(\tilde{f}_t) + \widetilde{s}\right) = \Phi\left(z_{it}(\tilde{f}_t) - \widetilde{s}\right) + \Phi\left(-z_{it}(\tilde{f}_t) - \widetilde{s}\right) > 0,$$

and it is easily seen that $h(z_{it}(\tilde{f}_t))$ is symmetric, namely $h(z_{it}(\tilde{f}_t)) = h(-z_{it}(\tilde{f}_t))$. Focusing on the non-negative values of $z_{it}(\tilde{f}_t)$ it is easily seen that

$$h(z_{it})) = \frac{z_{it}}{\sqrt{2\pi}}\left[e^{-0.5(z_{it}-\widetilde{s})^2} - e^{-0.5(z_{it}+\widetilde{s})^2}\right] > 0 \text{ for } \widetilde{s} > 0,$$

and by symmetry $h(z_{it})) \geq 0$, for all $\widetilde{s} \geq 0$. Hence, $q_{it} > 0$ for all i and t, and $\widetilde{s} \geq 0$. Therefore, it also follows that $\Psi'(f_t) > 0$, for all value of $w_{it} \geq 0$ and $s \geq 0$. Thus, by the fixed point theorem, $\Psi(f_t)$ must cut the horizontal axis but only once.

Proof of the consistency of $\tilde{f}_t$ as an estimator of f_t as $N \to \infty$.
Let

$$\begin{aligned} \Psi(f_t) &= \sum_{i=1}^{N} w_{it}\left\{z_{it}(f_t) + z_{it}(f_t)\left[\Phi\left(z_{it}(f_t) - \widetilde{s}\right) - \Phi\left(z_{it}(f_t) + \widetilde{s}\right)\right]\right. \\ &\qquad \left. + \delta^2\left[\phi\left(z_{it}(f_t) - \widetilde{s}\right) - \phi\left(z_{it}(f_t) + \widetilde{s}\right)\right]\right\} - a_{Nt}, \end{aligned}$$

and note that

$$\Psi(f_t) = -\sum_{i=1}^{N} w_{it}\eta_{it}.$$

Consider now the mean-value expansion of $\Psi(f_t)$ around $\tilde{f}_t$

$$\Psi(f_t) - \Psi(\tilde{f}_t) = \Psi'(\bar{f}_t)(f_t - \tilde{f}_t),$$

[1] Recall that the weights, w_{it}, are non-zero pre-determined constants, and in particular do not depend on f_t.

where $\bar{f}_t$ lies on the line segment between f_t and $\tilde{f}_t$. Since $\Psi(\tilde{f}_t) = 0$ and $\Psi'(\bar{f}_t) > 0$ for all $\bar{f}_t$ (as established above) we have

$$\tilde{f}_t - f_t = \frac{-\sum_{i=1}^{N} w_{it}\tilde{\eta}_{it}}{\Psi'(\bar{f}_t)}.$$

Recall that $\tilde{\eta}_{it} = \left(\sigma_s^2 + \sigma_\xi^2\right)^{-1/2} \left[\Delta p_{it} - E\left(\Delta p_{it} \left|\mathbf{h}_{it}\right.\right)\right]$, where $\mathbf{h}_{it} = (f_t, \mathbf{x}_{it}, p_{i,t-1})$, and hence $E\left(\tilde{\eta}_{it}\right) = 0$.. Further, conditional on f_t and $\mathbf{x}_{it}$, price changes, Δp_{it}, being functions of independent shocks v_i and ε_{it} over i, will be cross sectionally independent. Therefore, η_{it} will also be cross sectionally independent; although they need not be identically distributed even if the underlying shocks, v_i and ε_{it}, are identically distributed over i.

Given the above results we now have (for each t and as $N \to \infty$)

$$\left(\sum_{i=1}^{N} w_{it}^2\right)^{-1/2} \left(\tilde{f}_t - f_t\right) \backsim N\left(0, \vartheta_{\tilde{f}}^2\right),$$

where

$$\vartheta_{\tilde{f}}^2 = \lim_{N\to\infty} \left\{ \frac{\left(\sum_{i=1}^{N} w_{it}^2\right)^{-1} \sum_{i=1}^{N} w_{it}^2 Var(\tilde{\eta}_{it})}{\left[\Psi'(f_t)\right]^2} \right\}.$$

Note that as $N \to \infty$, $\sum_{i=1}^{N} w_{it}\tilde{\eta}_{it} \xrightarrow{p} 0$, and hence $\tilde{f}_t \xrightarrow{p} f_t$, since $\Psi'(f_t) > 0$ for all f_t.It must also be that $\bar{f}_t \xrightarrow{p} f_t$.

In the case where $w_{it} = 1/N$, we have

$$\vartheta_{\tilde{f}}^2 = \lim_{N\to\infty} \left\{ \frac{N^{-1} \sum_{i=1}^{N} Var(\tilde{\eta}_{it})}{\left[\Psi'(f_t)\right]^2} \right\}.$$

It also follows that $\tilde{f}_t - f_t = O_p\left(N^{-1/2}\right)$.. ■

B: Monte Carlo Simulations

We generated the log price series according to the baseline model, (**??**), and simulating the common factors as [a] first order autoregressive process. In our reference case, the sample size is set at $N = 50$, $T = 50$. In Table B.1, we report the average (across 1000 replications) of the point estimates of s, σ_ε, σ_s and σ_v and their average standard errors in different setups. Concerning the estimation of f_t, we compute the RMSE with respect to the true f_t and compare the standard deviation of the true f_t with that of the estimated f_t. Initial values for the estimation of f_t are set to $\overline{p}_t$. The standard errors of the parameter estimates are computed from the second derivatives of the full log-likelihood function. This table also reports the value of c computed from the point estimates of s, σ_ε and $\sigma_\omega = \sigma_f \sqrt{1-\rho^2}$, where σ_f is the standard deviation of the estimated f_t and ρ is the true autoregressive coefficient of the AR(1) process assumed for f_t.

Results reported in Table B.1 allow a comparison of the following cases: (i) with and without random effects, (ii) panels with N small, $N = 25$ versus $N = 50$, (iii) cases where the average frequency of price changes is 0.27 versus 0.12, (iv) the case of a small idiosyncratic component and a large common factor versus the case of a large idiosyncratic component and a relatively small common component, which corresponds to parameter values close to the estimates based on observed data. In general, estimated parameters are close to their true values. Our simulations show that the range of inaction is estimated with high precision. The estimate of the variance of the idiosyncratic component is closer to its theoretical value in the model with random effects. This drives the c above its true value, as it is related to the ratio of s to the size of the idiosyncratic and common shocks. Not surprisingly, the estimation of the common factor improves as the cross-section dimension increases. The results in Table B.1 also suggest that the estimation of f_t deteriorates as the frequency of price changes and the size of the common shock diminishes.

Our second set of Monte Carlo simulations consider the case of serial correlation of the idiosyncratic component. We model it as an AR(1) process where the variance of ε_{it} is identical to that of the base case. The results indicate that serial correlation in the idiosyncratic component introduces an upward bias in the estimated $\widehat{s}$ and $\widehat{\sigma}_s$.and a small downward bias in the estimates of $\widehat{\sigma}_\varepsilon$. The results are summarized in table B.2. The bias is negligible for low values of the serial correlation coefficient. It remains limited for small values of ρ (for $\rho = 0.50$, the estimate of s is only 0.03 higher than the true value). The bias becomes important only as serial correlation approaches the unit root case. However, because our measure of intrinsic price rigidity c is a function of $s/\sqrt{\sigma_\varepsilon{}^2 + \sigma_\omega{}^2}$, its computed value involves an upward bias that increases with the degree of serial correlation of ε_{it}. For $\rho = 0.50$, the bias amounts to 0.08.

The third set of Monte Carlo simulations examines the case of cross-sectional dependence. Cross-sectional dependence may be motivated on two grounds. First, local competition may lead outlets to be influenced by their neighbor pricing policies. Evidence on this can be found in Pinske et al. (2002) for US wholesale gasoline markets. Second, outlets of the same chain may have a common pricing policy, when pricing decision are centralized. In order to investigate this, two alternative specifications are chosen. The first is a Spatial Moving Average Model. The second is factor error structure where the cross-section dependence is generated according to a finite number of unobserved common factors. We include 10 factors for the 50 outlets considered in the experiments. The results are summarized in Table B.3.

As is well known in the literature on the linear factor model (Stock and Watson, 1998, Pesaran and Tosetti, 2007, Pesaran, 2006), "weak" cross sectional dependence (in the sense defined in Pesaran and Tosetti, 2007) does not affect the consistency of the estimates of the common factors using cross section averages or principle component approaches. The Monte Carlo experiments suggest that this property also holds in the case of our non linear factor model. Whether this result holds more generally clearly require further investigation.

Table B.1 - Monte Carlo Simulations

Average frequency of price changes ~ 0.27 with random effects							
	s	σ_ε	σ_s	σ_v			c
True values	0.15	0.05	0.01	0.025			0.082
	$\widehat{s}$	$\widehat{\sigma}_\varepsilon$	$\widehat{\sigma}_s$	$\widehat{\sigma}_v$	$RMSE(\widehat{f}_t)$	$\frac{RMSE(\widehat{f}_t)}{RMSE(f_t)}$	$\widehat{c}$
$N = 50, T = 50$	0.151	0.049	0.011	0.027	0.00019	1.0018	0.096
	(0.0014)	(0.0011)	(0.0013)	(0.0030)			
Average frequency of price changes ~ 0.27 without random effect							
	s	σ_ε	σ_s	σ_v			c
True values	0.15	0.05	0.01	0			0.082
	$\widehat{s}$	$\widehat{\sigma}_\varepsilon$	$\widehat{\sigma}_s$		$RMSE(\widehat{f}_t)$	$\frac{RMSE(\widehat{f}_t)}{RMSE(f_t)}$	$\widehat{c}$
$N = 50, T = 50$	0.150	0.049	0.007		0.00014	1.0018	0.099
	(0.0013)	(0.0011)	(0.0013)				
$N = 25, T = 50$	0.150	0.048	0.006		0.00029	1.0052	0.099
	(0.0019)	(0.0015)	(0.0018)				
Average frequency of price changes ~ 0.12 with random effect - large common component							
	s	σ_ε	σ_s	σ_v			c
True values	0.300	0.050	0.100	0.025			0.329
	$\widehat{s}$	$\widehat{\sigma}_\varepsilon$	$\widehat{\sigma}_s$	$\widehat{\sigma}_v$	$RMSE(\widehat{f}_t)$	$\frac{RMSE(\widehat{f}_t)}{RMSE(f_t)}$	$\widehat{c}$
$N = 50, T = 50$	0.302	0.047	0.103	0.029	0.00049	1.0052	0.430
	(0.0071)	(0.0017)	(0.0055)	(0.0036)			
Average frequency of price changes ~ 0.12 with random effect - large common component							
	s	σ_ε	σ_s	σ_v			c
True values	0.300	0.100	0.125	0.250			0.260
	$\widehat{s}$	$\widehat{\sigma}_\varepsilon$	$\widehat{\sigma}_s$	$\widehat{\sigma}_v$	$RMSE(\widehat{f}_t)$	$\frac{RMSE(\widehat{f}_t)}{RMSE(f_t)}$	$\widehat{c}$
$N = 100, T = 100$	0.307	0.099	0.131	0.247	0.00593	1.1841	0.380
	(0.0108)	(0.0027)	(0.0080)	(0.0246)			

Notes: 1000 replications, estimation by full ML. The figures in brackets are standard errors. f_t is simulated as a first order autoregressive process with intercept equal to 0.05 and slope equal to 0.90. $\sigma_f = 1$, except in the last simulation exercise (large idiosyncratic component) where $\sigma_f = 0.00625$. s is the size of the price inaction band, σ_ε^2 is the variance of the idiosyncratic component, σ_s^2 is the variance of s_{it} the threshold parameter for price changes. $\widehat{c}$ is estimated as $\widehat{s}^2/(\widehat{\sigma}\sqrt{6})$, where $\widehat{s}$ is the estimated size of the price inaction band, $\widehat{\sigma} = \sqrt{\widehat{\sigma}_\varepsilon^2 + \widehat{\sigma}_\omega^2}$, $\widehat{\sigma}_\varepsilon$ is the estimated standard deviation of the idiosyncratic component, and $\widehat{\sigma}_\omega$ is the estimated standard deviation of the common shock. $\frac{RMSE(\widehat{f}_t)}{RMSE(f_t)}$ is the ratio of the standard deviation of $\widehat{f}_t$ over the standard deviation of the true f_t.

Table B.2 - Monte Carlo Simulations with serially correlated idiosyncratic component

	s	σ^2_ε	σ^2_s			c
True values	0.35	0.005625	0.010			0.54
	$\widehat{s}$	$\widehat{\sigma^2_\varepsilon}$	$\widehat{\sigma^2_s}$	$RMSE(\widehat{f}_t)$	$\frac{RMSE(\widehat{f}_t)}{RMSE(f_t)}$	$\widehat{c}$
ρ=0	0.357 (0.020)	0.0038 (0.005)	0.0011 (0.002)	0.0021	1.343	0.55
ρ=0.10	0.359 (0.021)	0.007 (0.0004)	0.011 (0.002)	0.0021	1.356	0.56
ρ=0.50	0.379 (0.024)	0.0033 (0.0004)	0.013 (0.003)	0.0024	1.400	0.63
ρ=0.90	0.464 (0.042)	0.0022 (0.0004)	0.023 (0.006)	0.0030	1.425	1.00
ρ=0.95	0.510 (0.054)	0.0017 (0.0003)	0.029 (0.009)	0.0031	1.376	1.28
ρ=0.99	0.574 (0.087)	0.0012 (0.0003)	0.038 (0.015)	0.0029	1.162	2.00

Notes: 1000 replications, $N = 50$, $T = 50$, estimation by full ML. The figures in bracket are standard errors. f_t is simulated as a first order autoregressive process with intercept equal to 0.05 and slope equal to 0.75. ε_{it} is simulated as a first order autoregressive process with zero intercept and the serial correlation coefficient given by ρ. $\sigma_f= 0.057$ and $\sigma_\varepsilon= 0.075$. See also the notes to Table B.1.

Table B.3 - Monte Carlo Simulations with cross sectionally dependent idiosyncratic component

	s	σ^2_ε	σ^2_s			c
True values	0.35	0.005625	0.010			0.54
	$\widehat{s}$	$\widehat{\sigma^2_\varepsilon}$	$\widehat{\sigma^2_s}$	$RMSE(\widehat{f}_t)$	$\frac{RMSE(\widehat{f}_t)}{RMSE(f_t)}$	$\widehat{c}$
no cross-sectional dependence	0.357 (0.020)	0.0038 (0.0005)	0.011 (0.011)	0.0021	1.343	0.55
SMA(1)	0.357 (0.020)	0.0035 (0.0004)	0.011 (0.002)	0.0024	1.369	0.55
10 factors(2)	0.357 (0.020)	0.0036 (0.0004)	0.011 (0.002)	0.0024	1.375	0.55

Notes: Simulations are based on 1000 replications with $N = 50$ and $T = 50$. Estimation is by full ML. The figures in bracket are standard errors. f_t is simulated as $f_t= 0.05 + 0.75f_{t-1}+\omega_t$, ω_t~iid $N(0, \sigma^2_\omega)$, with $\sigma^2_\omega= 0.0002734$. See also the notes to Table B.1

(1) stands for the Spatial Moving Average model $\varepsilon_{it}= x_{it}+x_{i-1,t}+x_{i+1,t}$ with x_{it}~iid $N(0, \sigma^2_x)$.. The value of σ_x is set so that $\sigma_\varepsilon= 0.075$, the same value used in the experiments summarized in Table B.2.

(2) stands for the multifactor error structure $\varepsilon_{it}= \sum_{j=1}^{10} \gamma_{ij} z_{jt}+x_{it}$, where z_{jt}~iid $N(0, \sigma^2_j)$ and x_{it}~iid $N(0, \sigma^2_x)$ are drawn independently, with $\sigma^2_j= \sigma^2_x= 0.0028125$, $\gamma_{i1}= 1$ for $i = 1, ..., 5$, and 0 otherwise, $\gamma_{i2}= 1$ for $i = 6, ..., 10$, and 0 otherwise, $\gamma_{i3}= 1$, for $i = 11, ..., 15$, and so on.

C: Data Sources

The Belgian CPI data set contains monthly individual price reports collected by the Belgian National Statistical Institute (NSI) for the computation of the Belgian National and Harmonized Index of Consumer Prices. In its complete version, it covers the 1989:01 - 2005:12 period and contains more than 20,000,000 price records. For this project, we restricted the analysis to the product categories included in the Belgian CPI basket for the base year 1996, and restricted our sample period to the 1994:07 - 2003:02 period. Our data set covers only the product categories for which the prices are recorded throughout the entire year in a decentralized way, i.e. 65.5%. of the Belgian CPI basket for the base year 1996. The remaining 34.5% relate to product categories that are monitored centrally, such as housing rents, electricity, gas, telecommunications, health care, newspapers and insurance services and to seasonal product categories. Price reports take into account all types of rebates and promotions, except those relating to the winter and summer sales period, which typically take place in January and July. In addition to the price records, the Belgian CPI data sets provides information on the location of the seller, a seller identifier, the packaging of the product and the brand of the product. The price concept use in this article is the price per unit.

The French CPI data set contains more than 13,000,000 monthly individual price records collected by the INSEE for the computation of the French National and Harmonized Index of Consumer Prices. It covers the period July 1994:07 - February 2003. This data set covers 65.5%. of the French CPI basket. Indeed, the prices of some categories of goods and services are not available in our sample: centrally collected prices - of which major items are car prices and administered or public utility prices (e.g. electricity)- as well as other types of products such as fresh food and rents. At the COICOP 5-digit level, we have access to 128 product categories out of 160 in the CPI. As a result, the coverage rate is above 70% for food and non-energy industrial goods, but closer to 50% in the services, since a large part of services prices are centrally collected, e.g. for transport or administrative or financial services. Each individual price quote consists of the exact price level of a precisely defined product. What is meant by "product" is a particular product, of a particular brand and quality, sold in a particular outlet. The individual product identification number allows us to follow the price of a product through time, and to recover information on the type of outlet (hypermarket, supermarket, department store, specialized store, corner shop, service shop, etc.), the category of product and the regional area where the outlet is located (for confidentiality reasons, a more precise location of outlets was not made available to us). The sequences of records corresponding to such defined individual products are referred to as price trajectories. Importantly, if in a given outlet a given product is definitively replaced by a similar product of another brand or of a different quality, a new identification number is created, and a new price trajectory is started. On top of the above mentioned information, the following additional information is recorded: the year and month of the record, a qualitative "type of record" code and (when relevant) the quantity sold. When relevant, division by the indicator of the quantity is used in order to recover a consistent price per unit.

Confidentiality data restrictions : We are not allowed to provide anyone with the micro price reports underlying this work. However, a data set containing simulated data and the MatLab or SAS codes of the estimation procedures are available on request.