Forecasting and Predictive Analytics

Answers to Homework 2

November 26, 2017

Forecasting Prices

This project aims at performing a first study of price dynamics. The first section asks you to analyze so-called *Phillips curve* models. The second part focuses on an empirical application to forecasting some retail prices for US markets.

This study goes beyond economics as it is an introduction to the issue of forecasting when you need to take into the forecasts agents themselves make. In practice the latter forecasts are often observables, either via surveys or using information drawn from futures/forward financial markets (such as options, forward contracts...).

1 Price and cost dynamics

Consider a firm that needs to sign a contract and for this reason tries to estimate what the prices its competitors are going to set. For this we assume that the market is very competitive so there is a "representative firm" whose behavior is the average of all (think here that it could e.g. be an agricultural firm). The price P_t this representative firm sets is assumed to be function of its current cost of labor and energy C_t. The firm also wants assess its ability to sustain the price it sets: it does not want to renegotiate the contract to adjust its price, hence the current price depends on its forecast of future prices of competitors and, if we assume all firms are similar (that's the notion of "representative firm"), the forecast of its own future price, which we write $P^e_{t+1|t}$.

Denote $p_t = \log P_t$ (and similarly for other variables) then if the firm minimizes the cost of adjusting its price over its forecast horizon, it can be shown that the price follows the "forward looking" law of motion:

$$\Delta p_t = \beta \left(p^e_{t+1|t} - p_t \right) + \lambda c_t$$

where $\beta \in (0, 1)$ and which can be written in terms of price inflation $\pi_t = \Delta p_t$:

$$\pi_t = \beta \pi^e_{t+1|t} + \lambda c_t. \tag{1}$$

We consider first the case where firms know the law of motion of c_t and compute their forecasts using $E_t[\cdot] = E[\cdot|\mathcal{I}_t]$, the expectation conditional upon $\mathcal{I}_t$, the information available at time t. We assume throughout that the information set consists of the history of (π_t, c_t). Somehow, this assumes that agents also know the law of motion (data generating process) of π_t. Indeed, (1) then becomes

$$\pi_t = \beta E_t \pi_{t+1} + \lambda c_t. \tag{2}$$

We first solve this expression forward, for any $K > 0$,

$$\begin{aligned}\pi_t &= \beta E_t[\pi_{t+1}] + \lambda c_t = \beta E_t[\beta E_{t+1}[\pi_{t+2}] + \lambda c_{t+1}] + \lambda c_t \\ &= ... = \beta^{K+1} E_t ... E_{t+K} \pi_{t+K+1} + \lambda c_t + \lambda \sum_{i=1}^{K} E_t ... E_{t+i-1}\left[\beta^i c_{t+i}\right].\end{aligned}$$

If the $\beta \in (0,1)$ and so-called *transversality condition*[1] holds:

$$\beta^{K+1} E_t ... E_{t+K}[\pi_{t+K+1}] \underset{K\to\infty}{\to} 0. \tag{3}$$

Using the Law of Iterated Expectations to simplify the previous expressions:

$$E_t[E_{t+1}[c_{t+1}]] = E_t[c_{t+1}]$$

so

$$\pi_t = \lambda c_t + \lambda \sum_{i=1}^{\infty} E_t\left[\beta^i c_{t+i}\right].$$

The expression $\sum_{i=1}^{\infty} E_t\left[\beta^i c_{t+i}\right]$ is called expected present-value of future costs discounted at fixed rate β.

1. Assume the firm is thinking of using a technology which will reduce its cost by δ in k periods (such as shifting to solar electricity). If this technology adoption is known and announced at time t, what is the impact on inflation at periods $t+j$ for $j = 0,...,k$?

   ANSWER: There are ways to interpret this question, the most likely is that from $t+k$ onwards, costs are reduced by a permanent amount δ, so for $i \geq 0$, c_{t+k+i} becomes $c^*_{t+k+i} = c_{t+k+i} - \delta$. Then π_{t+j} becomes π^*_{t+j} where, letting $1_{\{\mathcal{H}\}}$ the indicator function that takes value 1 if $\mathcal{H}$ is true and 0 otherwise:

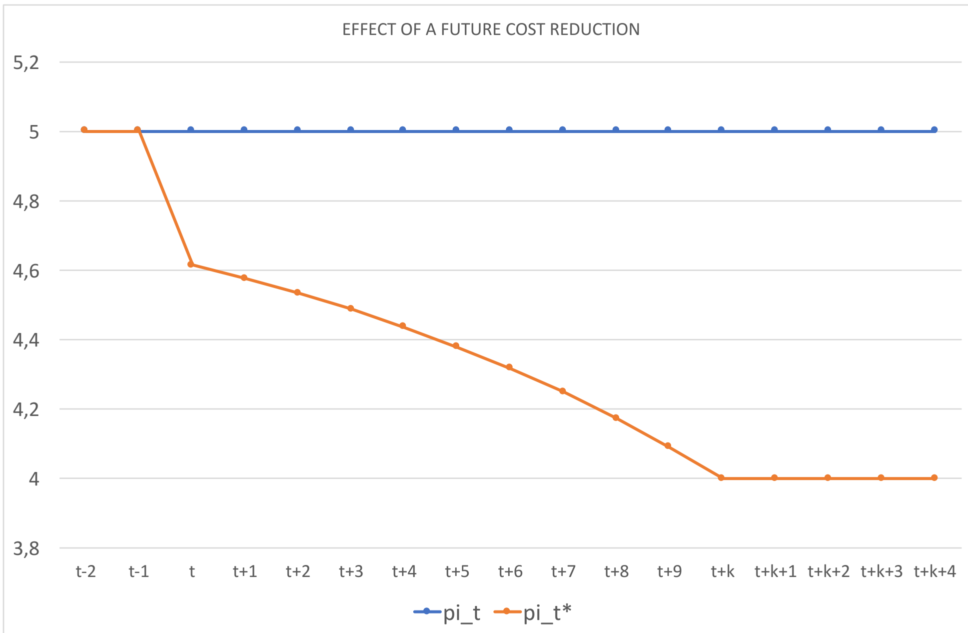
$$\begin{aligned}\pi^*_{t+j} &= \lambda\left(c_{t+j} - \delta \times 1_{\{t+j \geq t+k\}}\right) + \lambda \sum_{i=1}^{\infty} E_t\left[\beta^i\left(c_{t+j+i} - \delta \times 1_{\{t+j+i \geq t+k\}}\right)\right] \\ &= \pi_{t+j} - \lambda\delta \times 1_{\{t+j \geq t+k\}} - \lambda \sum_{i=1}^{\infty} \beta^i \delta \times 1_{\{t+j+i \geq t+k\}} \\ &= \pi_{t+j} - \lambda\delta \times 1_{\{j \geq k\}} - \lambda\delta \sum_{i=\max(k-j,1)}^{\infty} \beta^i \\ &= \pi_{t+j} - \lambda\delta \frac{\beta^{\max(k-j,0)}}{1-\beta}\end{aligned}$$

[1] Also called non-Ponzi scheme in finance, you may want to have a look at who Ponzi was.

i.e.

$$\pi^*_{t+j} = \pi_{t+j} - \begin{cases} \frac{\lambda\delta}{1-\beta}, & \text{if } j \geq k, \\ \frac{\lambda\delta\beta^{(k-j)}}{1-\beta}, & \text{if } j < k. \end{cases}$$

so the reduction in cost is progressively "priced in" in that the decrease is progressive, as in the following graph where π_t is obtained assuming the cost c_t is constant throughout.



2. Assume, in this question only, that c_t follows a random walk: $c_t = c_{t-1} + \varepsilon_t$ where $\varepsilon_t \overset{i.i.d}{\sim} \mathsf{N}\left(0, \sigma^2_\varepsilon\right)$.

   (a) What is the value of $E_t\left[c_{t+i}\right]$ for $i \geq 0$?

   ANSWER: $E_t\left[c_{t+i}\right] = c_t$

   (b) Derive a relation between π_t and c_t. Does the condition (3) hold?

   ANSWER:

$$c_t = \lambda c_t + \lambda \sum_{i=1}^{\infty} \beta^i c_t = \lambda c_t \sum_{i=0}^{\infty} \beta^i = \frac{\lambda}{1-\beta} c_t.$$

   Now

$$E_t\left[\beta^k c_{t+k}\right] = \beta^k c_t \underset{k\to\infty}{\to} 0$$

   so the transversality (or non-Ponzi) condition holds.

(c) The previous value π_t is called the fundamental value. Now, consider $\pi_t^* = \pi_t + B_t$. Show that if

$$B_t = \beta E_t [B_{t+1}] \tag{4}$$

then π_t^* is also solution to (2), i.e. $\pi_t^* = \beta E_t \pi_{t+1}^* + \lambda c_t$. Show that it is possible to assume that B_t follows an autoregressive process of order 1. Is it, then, stationary?

ANSWER:

$$\pi_t^* = \pi_t + B_t = \beta E_t [\pi_{t+1} + B_{t+1}] + \lambda c_t = \beta E_t \left[\pi_{t+1}^*\right] + \lambda c_t$$

Now let $\zeta_{t+1} = B_{t+1} - E_t (B_{t+1})$ so $E_t (\zeta_{t+1}) = 0$ (ζ_t is called a martingale difference sequence), it follows, by definition that

$$B_{t+1} = E_t (B_{t+1}) + \zeta_{t+1}$$

with $E_t (B_{t+1}) = \frac{1}{\beta} B_t$ so

$$B_{t+1} = \frac{1}{\beta} B_t + \zeta_{t+1}$$

i.e. B_t follows an AR(1) with coefficient $\beta^{-1} > 1$, i.e. nonstationary and explosive.

(d) Blanchard and Watson (1982) suggested a solution of the form

$$B_{t+1} = \begin{cases} \frac{1}{\beta\pi} B_t + \zeta_{t+1} & \text{with probability } \pi; \\ \zeta_{t+1} & \text{with probability } 1 - \pi; \end{cases}$$

with $E_t \zeta_{t+1} = 0$.

Show that this definition is such that π_t^* satisfies (2). Argue why this is called a *rational bubble* (you might want to simulate one example in Excel or R to convince yourselves). What type of dynamics does this imply for inflation π_t^*? Explain why hyperfinflation is a phenomenon that can be sustained if every firms believes it will continue.[2] Can you find a way to stop it once it has started?

In the remainder of this project, we will assume that price inflation is the fundamental value and that no bubble is observed.

ANSWER: From the expression above

$$\begin{aligned} E_t [B_{t+1}] &= \pi \cdot E_t \left[\frac{1}{\beta\pi} B_t + \zeta_{t+1}\right] + (1 - \pi) \cdot E_t [\zeta_{t+1}] \\ &= \pi \cdot E_t \left[\frac{1}{\beta\pi} B_t\right] + E_t [\zeta_{t+1}] \\ &= \frac{1}{\beta} B_t \end{aligned}$$

[2] See also the historical examples

https://en.wikipedia.org/wiki/Tulip_mania

https://en.wikipedia.org/wiki/Mississippi_Company

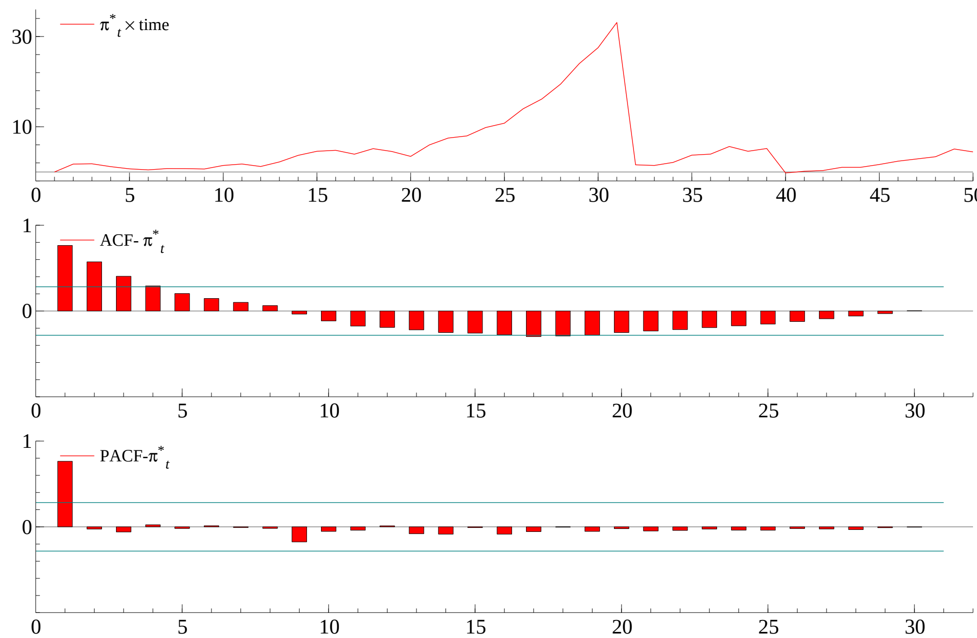
as well as the articles on the website. We may wonder whether you should be careful with bitcoins.

hence

$$\beta E_t\left[B_{t+1}\right] = B_t$$

if B_t belongs to the information set available at time t.

Now from the definition of B_{t+1}, it either follow an AR(1) with coefficient $1/(\beta\pi) > 1$ (i.e. an explosive process), or a white noise. While in a succession of probability π events, the bubble rides on: the value π^*_{t+j} grows exponentially fast, and when a $(1-\pi)$-probability event happens, the bubble bursts, as in the following graph.



3. We now assume that the firm does not know the DGP of its costs and hence use so-called "adaptive expectations"

$$\pi^e_{t+1|t} = \pi^e_{t|t-1} + \alpha\left(\pi_t - \pi^e_{t|t-1}\right). \tag{5}$$

   (a) Explain what this method amounts to?

   ANSWER: this method implies that $\pi^e_{t+1|t}$ is the value that is obtained from applying the exponential smoother to π_t :

$$\pi^e_{t+1|t} = \alpha\sum_{j=0}^{\infty}(1-\alpha)^j\,\pi_{t-j}.$$

   (b) Show that Equations (1) and (5) imply that

$$\pi_t = \frac{1-\alpha}{1-\beta\alpha}\pi_{t-1} + \frac{\lambda}{1-\beta\alpha}c_t - \frac{\lambda(1-\alpha)}{1-\beta\alpha}c_{t-1}$$

   Assume that c_t follows an AR(1) : $c_t = (1-\alpha)\,c_{t-1} + u_t$ where $u_t \sim \mathsf{NID}\left(0, \sigma_u^2\right)$. Contrast the dynamics where forecasts are based on the true mathematical expectation vs. adaptive expectations.

ANSWER: The system

$$\pi_t = \beta \pi^e_{t+1|t} + \lambda c_t$$
$$\pi^e_{t+1|t} = (1-\alpha)\,\pi^e_{t|t-1} + \alpha \pi_t$$

implies that

$$\begin{aligned}\pi_t - (1-\alpha)\,\pi_{t-1} &= \beta\left[\pi^e_{t+1|t} - (1-\alpha)\,\pi^e_{t|t-1}\right] + \lambda\left[c_t - (1-\alpha)\,c_{t-1}\right] \\ &= \beta\left[\alpha \pi_t\right] + \lambda\left[c_t - (1-\alpha)\,c_{t-1}\right].\end{aligned}$$

Hence

$$(1-\beta\alpha)\,\pi_t = (1-\alpha)\,\pi_{t-1} + \lambda\left[c_t - (1-\alpha)\,c_{t-1}\right]$$

which yields the required result when dividing both sides by $(1-\beta\alpha)$. Now assume c_t follows an AR(1) with coefficient $(1-\alpha)$ so $c_t - (1-\alpha)\,c_{t-1} = u_t$. Hence, under Adaptive Expectations (AE), π_t also follows an AR(1) :

$$\pi_t = \frac{1-\alpha}{1-\beta\alpha}\pi_{t-1} + \frac{\lambda}{1-\beta\alpha}u_t.$$

Here, under adaptive expectations inflation follows a stationary (since $\left|\frac{1-\alpha}{1-\beta\alpha}\right| < 1$) AR(1) process. By contrast, under rational expectations

$$\pi_t = \lambda c_t + \lambda \sum_{i=1}^{\infty} E_t\left[\beta^i c_{t+i}\right]$$

where $E_t\left[c_{t+i}\right] = E_t\left[(1-\alpha)^i c_t + \sum_{j=1}^{i}(1-\alpha)^{i-j} u_{t+j}\right] = (1-\alpha)^i c_t$. Hence

$$\pi_t = \lambda c_t \sum_{i=0}^{\infty}\left(\beta(1-\alpha)\right)^i = \frac{\lambda}{1-\beta(1-\alpha)}c_t,$$

i.e., since c_t follows an AR(1), under Rational Expectations (RE):

$$\pi_t = (1-\alpha)\,\pi_{t-1} + \frac{\lambda}{1-\beta(1-\alpha)}u_t.$$

Contrasting AE and RE, we notice that under both assumptions π_t follows an AR(1) driven by the innovations u_t. Under AE, the autoregressive coefficient is $\frac{1-\alpha}{1-\beta\alpha}$, i.e. greater than $1-\alpha$ so AE induces more persistence (higher autocorrelation). Now, regarding the variance of the innovations of the autoregressive process, $\frac{\lambda}{1-\beta\alpha}$ can be smaller or larger than $\frac{\lambda}{1-\beta(1-\alpha)}$ depending on whether α is greater than $1-\alpha$ or not.

4. We now consider the situation where costs follow the process

$$c_t = x_t + v_t$$
$$x_t = x_{t-1} + u_t$$

where (u_t, v_t) are Gaussian white noise and independent of each other. c_t is then the sum of an unobserved (latent) random walk and of a white noise. This is the case for instance when x_t is not very accurately measured.

(a) Compute (analytically) the autocorrelation function of $\Delta c_t = c_t - c_{t-1}$ (note: this should be easy).

ANSWER: $\Delta c_t = \Delta x_t + \Delta v_t$ where $\Delta x_t = u_t$ so

$$\Delta c_t = u_t + \Delta v_t$$

with variance $\mathsf{V}(\Delta c_t) = \mathsf{V}(u_t + \Delta v_t) = \sigma_u^2 + 2\sigma_v^2$ and covariance

$$\mathsf{Cov}(\Delta c_t, \Delta c_{t-j}) = \begin{cases} -\sigma_v^2, & \text{if } j = 1, \\ 0 & \text{if } j > 1. \end{cases}$$

so

$$\mathsf{Corr}(\Delta c_t, \Delta c_{t-1}) = -\frac{1}{1 + 2\sigma_v^2/\sigma_u^2}$$

and $\mathsf{Corr}(\Delta c_t, \Delta c_{t-j}) = 0$ for $j > 1$. Notice that under this process c_t is a sum of a random walk (x_t) and an additional white noise v_t.

(b) Show that the Wold decomposition theorem implies that there exists a white noise process ε_t such that

$$\Delta c_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

for some θ. What is the value of θ as a function of σ_u^2 and σ_v^2, the variances of u_t and of v_t.

ANSWER: Since $\Delta c_t = u_t + \Delta v_t$, Δc_t is weakly stationary and the Wold decomposition theorem states that it admits an MA(∞) representation

$$\Delta c_t = \varepsilon_t + \sum_{j=1}^{\infty} \theta_j \varepsilon_{t-j}.$$

Now if any $\theta_j \neq 0$, then $\mathsf{Cov}(\Delta c_t, \Delta c_{t-j}) \neq 0$ since it involves the term $\theta_j \sigma_\varepsilon^2$. Hence from what we saw above, the only nonzero θ_j is θ_1. It follows that Δc_t follows an MA(1), $\Delta c_t = \varepsilon_t + \theta \varepsilon_{t-1}$. Now

$$\mathsf{V}(\Delta c_t) = \left(1 + \theta^2\right) \sigma_\varepsilon^2$$
$$\mathsf{Cov}(\Delta c_t, \Delta c_{t-j}) = \theta \sigma_\varepsilon^2$$

so we can solve

$$\mathsf{Corr}(\Delta c_t, \Delta c_{t-1}) = \frac{\theta}{1 + \theta^2} = -\frac{1}{1 + 2\sigma_v^2/\sigma_u^2}$$

so $1+\left(1+2\sigma_v^2/\sigma_u^2\right)\theta+\theta^2=0$, i.e.,

$$\theta=\frac{-\left(1+2\sigma_v^2/\sigma_u^2\right)+\sqrt{\left(1+2\sigma_v^2/\sigma_u^2\right)^2-4}}{2}$$

(the other solution does not ensure that the MA(1) is invertible).

(c) What is the resulting dynamics for π_t assuming that the information set consists of the history of (π_t, c_t)?

ANSWER: What we need to understand is which of the two representations above we should use, indeed if x_t is observable then $E_t c_{t+j} = E_t x_{t+j} = x_t$ but if the information set consists of (π_t, c_t) we cannot recover x_t (or alternatively, u_t), we can only observe ε_t (by inverting the MA lag polynomial $1+\theta L$). Hence,

$$E_t\left[\Delta c_{t+j}\right]=E\left[\Delta c_{t+j}|\Delta c_t,\pi_t,\Delta c_{t-1},\pi_{t-1},...\right]=E\left[\Delta c_{t+j}|\varepsilon_t,\pi_t,\varepsilon_{t-1},\pi_{t-1},...\right]$$

i.e.,

$$E_t\left[\Delta c_{t+j}\right]=E_t\left[\varepsilon_{t+j}+\theta\varepsilon_{t+j-1}\right]=\begin{cases}\theta\varepsilon_t, & \text{if } j=1\\ 0, & \text{if } j\geq 2\end{cases}$$

and, for $j\geq 1$,

$$E_t\left[c_{t+j}\right]=E_t\left[c_t+\sum_{k=1}^{j}\Delta c_{t+k}\right]=c_t+\theta\varepsilon_t.$$

Now regarding the resulting dynamics for π_t,

$$\begin{aligned}\pi_t &= \lambda c_t+\lambda\sum_{i=1}^{\infty}E_t\left[\beta^i c_{t+i}\right]\\ &= \frac{\lambda}{1-\beta}c_t+\frac{\lambda\theta\beta}{1-\beta}\varepsilon_t.\end{aligned}$$

so $\pi_t-\frac{\lambda\theta\beta}{1-\beta}\varepsilon_t=\frac{\lambda}{1-\beta}c_t=\frac{\lambda}{1-\beta}\left[c_{t-1}+\varepsilon_t+\theta\varepsilon_{t-1}\right]$,

$$\begin{aligned}\pi_t-\frac{\lambda\theta\beta}{1-\beta}\varepsilon_t &= \pi_{t-1}-\frac{\lambda\theta\beta}{1-\beta}\varepsilon_{t-1}+\frac{\lambda}{1-\beta}\left[\varepsilon_t+\theta\varepsilon_{t-1}\right]\\ \pi_t &= \pi_{t-1}+\frac{\lambda\left(1+\theta\beta\right)}{1-\beta}\varepsilon_t+\frac{\lambda\theta\left(1-\beta\right)}{1-\beta}\varepsilon_{t-1}\end{aligned}$$

i.e. both c_t and π_t follow ARIMA$(0,1,1)$ processes.

2 Empirical Analysis

SEE ATTACHED FILE