# Forecasting and Predictive Analytics

Homework 2

November 13, 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 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 $\pi_t = \Delta p_t$ :

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

**Remark 1** *Note that Equation (1) is also the present value model for stock returns where, instead of $c_t$, the so-called "forcing variable" is expected cash flow/dividends: the price of an asset at time t, $P_t$, depends on the cash flow $D_t$ (e.g. dividend which is announced at time t) it generates between t and $t+1$ as well as forecasts of future prices.*

*The reason behind this model is that an investor who wishes to invest $P_t$ at time t has the choice between (i) investing in the firm or (ii) put the $P_t$ in a bank as a safe deposit. If the investor follows strategy (i), she receives in period $t+1$ the cash flow $D_t$ (i.e. if she buys stocks, the cash flow is the dividend) and she can resell the investment at price $P_{t+1}$. Given that she does not know at time t the values of $P_{t+1}$, her expected gain is at $t$ : $P^e_{t+1|t} + D_t$. If, by contrast the investor decides to deposit $P_t$, she earns the safe return $1+R$ so she knows that she possesses $(1+R)P_t$ in period $t+1$.*

*Since we observe both types of investments at each point in time, it means that neither provides a certain higher return, or else every investor would choose the investment with certain higher return. This is called the no-arbitrage condition and implies that*

$$(1+R)P_t = P^e_{t+1|t} + D_t.$$

*This yields*[1]

$$P_t = \frac{1}{1+R}P^e_{t+1|t} + \frac{1}{1+R}D_t$$

*which looks like the dynamics generated by Equation (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 $(\pi_t, c_t)$.

Somehow, this assumes that agents also know the law of motion (data generating process) of $\pi_t$. 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 $K > 0$,

$$\begin{aligned}\pi_t &= \beta E_t[\pi_{t+1}] + \lambda c_t = \beta E_t[\beta E_{t+1}[\pi_{t+2}] + \lambda c_{t+1}] + \lambda c_t \\ &= ... = \beta^{K+1} E_t ... E_{t+K}\pi_{t+K+1} + \lambda \sum_{i=0}^{K} E_t ... E_{t+i}\left[\beta^i c_{t+i}\right].\end{aligned}$$

If the $\beta\lambda \in (0,1)$ and so-called *transversality condition*[2] holds:

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

[1] For those who are interested, you should have a look at the 2013 Nobel lecture of Rob Shiller, who essentially received his Nobel prize for his study of this equation.

https://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2013/shiller-lecture.html

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

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

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

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 $\beta$.

1. Assume the firm is thinking of using a technology which will reduce its cost by $\delta$ 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 $j = 0,...,k$?

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

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

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

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

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

   then $\pi_t^*$ 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?

   (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 $\pi_t^*$ 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 $\pi_t^*$? Explain why hyperfinflation is a phenomenon that can be sustained if every firms believes it will continue.[3] 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.

[3] 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.

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?

   (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.

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).

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

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

   for some $\theta$. What is the value of $\theta$ as a function of $\sigma_u^2$ and $\sigma_v^2$, the variances of $u_t$ and of $v_t$.

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

# 2 Empirical Analysis

The accompanying dataset contains monthly data (from January 1984) for the U.S. price of Ground Beef. Your ultimate aim is to produce forecasts for the last three months 2017 (i.e. 1 to 3 months ahead), you will have to provide the forecasts (on a sheet on the website) and I will assess them. For the purpose of the homework, you can work with any training and testing sets as you wish (these are often called in-sample forecasts and pseudo out-of-sample forecasts). $p_t$ and $\pi_t$ denote respectively the log of the price of ground beef and corresponding beef price inflation

1. Are these two variables stationary (using a simple method, e.g. looking at the time series plots and at the ACF)?

2. We now consider only the log price $p_t$ and want to forecast it.

   (a) Propose a method for forecasting $p_t$ based on trend decomposition/smoothing (exponential smoothing, Holt-Winters, polynomial trend....). Make sure to be clear in your assumptions.

   (b) Fit an ARMA model to $\pi_t = \Delta p_t$ and use it to forecast $p_t$ one to three steps ahead. Notice that you can use iterative or direct multi-step forecasts:
   **Iterative:** A model is built by estimating $y_t$ using $\mathcal{I}_{t-1}$ information. For example, in an AR(1),

   $$y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t.$$

   Multi-step ahead forecasts are computed by iterating forward $h$-steps. In the case of an AR(1), the $h$-step ahead forecast is $\widehat{y}_{t+h|t} = \sum_{i=1}^{h} \phi_1^{i-1}\phi_0 + \phi_1^h y_t$.
   **Direct forecast:** A model is built by estimating $y_t$ using $\mathcal{I}_{t-h}$ information. For example, in an AR(1)-like specification,

   $$y_t = \psi_0 + \psi_h y_{t-h} + \eta_t.$$

   Multi-step ahead forecasts are computed as one-step ahead forecasts from the direct model. In the case of an AR(1)-like specification, the $h$-step ahead forecast is $\widetilde{y}_{t+h|t} = \psi_0 + \psi_h y_t$. When $h = 1$, these two methods are identical.

3. Assess and compare your forecasts in your testing sample (including using Mincer-Zarnowitz regressions). Do not forget to produce also true out-of-sample forecasts for the end of 2017.

4. Are we in 2017 in a "Super-cycle" as claimed by the industry observer Simon Quilty at beef.com?
   https://www.beefcentral.com/trade/are-the-us-and-global-meat-markets-entering-into-a-super-demand-cycle-for-201718/

**Remark 2** *To forecast one-month ahead the monthly variable $y_t$ means obtaining forecasts of $y_{t+1}$ given information available at time $t$ (denoted $\mathcal{I}_t$). On page 90 of the handout, it is shown that the* optimal *forecast, i.e. the one which minimizes the mean square forecast error, is the conditional expectation. So we are going to use the conditional expectation as a forecast:*

$$y_{t+h|t} \overset{def}{=} \mathsf{E}\left[y_{t+h}|\mathcal{I}_t\right]$$

*so if your model is* $y_t \sim AR(p)$*:*

$$y_t = \tau + \alpha_1 y_{t-1} + ... \alpha_p y_{t-p} + \varepsilon_t$$

*then the so called* iterated forecasts *are:*

$$\begin{aligned}
y_{t+1|t} &= \widehat{\tau} + \widehat{\alpha}_1 y_t + ... + \widehat{\alpha}_p y_{t-p+1} \\
y_{t+2|t} &= \widehat{\tau} + \widehat{\alpha}_1 y_{t+1|t} + \widehat{\alpha}_2 y_t + ... + \widehat{\alpha}_p y_{t-p+2} \\
&\vdots \\
y_{t+p|t} &= \widehat{\tau} + \widehat{\alpha}_1 y_{t+p-1|t} + \widehat{\alpha}_2 y_{t+p-2|t} + ... + \widehat{\alpha}_p y_t \\
y_{t+p+1|t} &= \widehat{\tau} + \widehat{\alpha}_1 y_{t+p|t} + \widehat{\alpha}_2 y_{t+p-1|t} + ... + \widehat{\alpha}_p y_{t+1|t} \\
&\vdots
\end{aligned}$$

*You see clearly that because we are conditioning on* $\mathcal{I}_t$*, all the forecasts will end up being functions of* $y_t, ..., y_{t-p+1}$ *only. This means that when we have a specific forecast horizon in mind (e.g. 3-month) instead of having a model for* $y_t$ *as a function of* $\mathcal{I}_{t-1}$*, and then computing the forecasts forward, we can simply use a* direct forecast*, which consists in estimating an* $AR(p+2)$ *model where the first two coefficients are imposed to be zero, i.e. estimating the model:*

$$y_t = \gamma + \beta_3 y_{t-3} + \beta_4 y_{t-4} + ... \beta_{p-2} y_{t-p+2} + \eta_t$$

*and compute the forecast*

$$y_{t+3|t} = \widehat{\gamma} + \widehat{\beta}_3 y_t + \widehat{\beta}_4 y_{t-1} + ... \widehat{\beta}_{p-2} y_{t-p}.$$