EC508: Econometrics
Forecasting and Dynamic Causal Effects

Jean-Jacques Forneron

Spring, 2023

Boston University

Forecasting

- Given the data $(Y_1, \dots, Y_T, X_1, \dots, X_T)$ we can forecast Y in the next period using:

$$\begin{aligned}\hat{Y}_{T+1|T} =& \hat{\beta}_0 + \hat{\beta}_1 Y_T + \cdots + \hat{\beta}_p Y_{T-p+1} \\ & + \hat{\delta}_{1,1} X_{1,T} + \cdots + \hat{\delta}_{1,q_1} X_{1,T-q_1+1} \\ & + \cdots + \hat{\delta}_{k,q_k} X_{k,T-q_k+1}\end{aligned}$$

- When Y_{T+1} is released, we can compute the forecast error:

$$Y_{T+1} - \hat{Y}_{T+1|T} = X_T' [\underbrace{\beta - \hat{\beta}}_{\text{estimation error}}] + \underbrace{u_{T+1}}_{\text{fundamental uncertainty}}$$

Forecast Errors

- Forecast error:

$$Y_{T+1} - \hat{Y}_{T+1|T} = X_T'[\underbrace{\beta - \hat{\beta}}_{\text{estimation error}}] + \underbrace{u_{T+1}}_{\text{fundamental uncertainty}}$$

- Since $\hat{\beta}$ is estimated using data from the past and u_{t+1} comes from the future, we can split to mean squared forecast error (MSFE) into:

$$\mathbb{E}_T([Y_{T+1} - \hat{Y}_{T+1|T}]^2) = var(u_{T+1}) + X_T' var(\hat{\beta}) X_T$$

- More parameters: larger variance $var(\hat{\beta})$ (declines with T), lower variance $var(u_{T+1})$

$\Rightarrow$ tradeoff between estimation accuracy and the variance of the error term u_{T+1}

RMFSE; Pseudo Out-of-Sample Forecasting

- Root Mean Squared Forecast error (RMSFE): $\sqrt{MSFE}$
- Suppose we estimate $\hat{\beta}$ using data $t = 1, \ldots, T - t_1 - 1$ and compute forecasts for $t = T_1, \ldots, T$
- This allows us to replicate the forecasting conditions
- We can compute the approximation:

$$RMSFE \simeq \sqrt{\frac{1}{T - T_1 - 1} \sum_{t=T_1}^{T-1} (Y_{t+1} - \hat{Y}_{t+1|t})^2}$$

- This is called **pseudo out-of-sample forecasting**
- We can use these estimates to compare different models (p, q) to see which one(s) historically perform best out-of-sample

Forecast Intervals

- If $u_{t+1} \sim \mathcal{N}(0, \sigma_u^2)$; given that $\hat{\beta} - \beta$ is approximately normal then we can make the approximation

$$\hat{Y}_{T+1|T} - Y_{T+1} \sim \mathcal{N}(0, MSFE)$$

- A 95% forecast interval, which contains the true Y_{T+1} 95% of the time, is given by

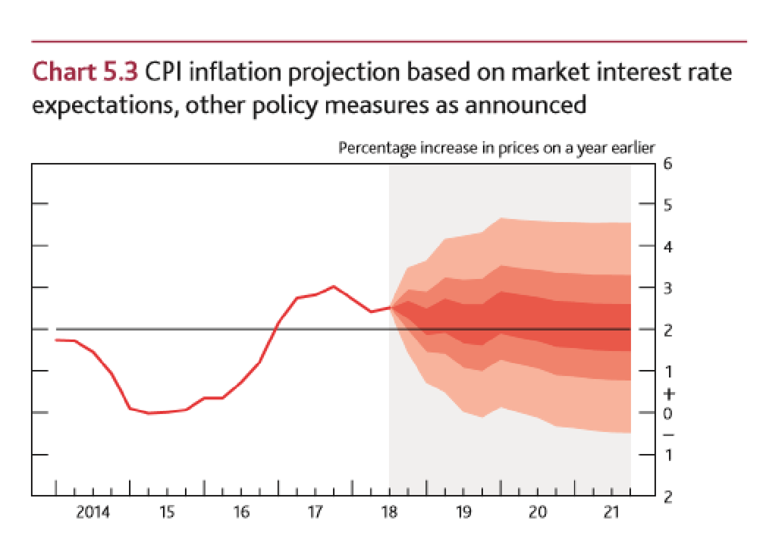
$$\{\hat{Y}_{T+1|T} \pm 1.96 \times RMSFE\}$$

- RMSFE is approximated with pseudo out-of-sample forecasts or with $\hat{\sigma}_u$ and $\hat{V}_\beta$ the asymptotic variance of $\hat{\beta}$

Forecast Intervals in Central Banks

- The Bank of England publishes its Inflation Reports with forecasts of future inflation along with confidence bands that summarize the uncertainty around its forecasts
- This is useful in communicating how it intends to achieve its objective of maintaining inflation close to its 2% target

Forecast Intervals in Central Banks



Distributed Lag Model and Dynamic Causal Effects

- The Distributed Lag (DL) model is

$$Y_t = \beta_0 + \beta_1 X_t + \cdots + \beta_q X_{t-q} + u_t$$

- If X_t is exogenous then the regression has a causal interpretation: e.g. effect of oil price on GDP/inflation
- Suppose we increase X_t by 1 and then revert to the mean (say 0) for ever afterwards
- Effect: at t - β_1, at $t+1$ - β_2,...
- The effect is persistent over time, a one-time event will affect Y_t now and into the future
- The cumulative multiplier is the sum of all the effects this one-time shock will have:

$$\beta_1 + \beta_2 + \cdots + \beta_q$$

Computing Cumulative Multipliers

- The Distributed Lag (DL) model is

$$Y_t = \beta_0 + \beta_1 X_t + \cdots + \beta_q X_{t-q} + u_t$$

- To compute $\beta_1 + \beta_2 + \cdots + \beta_q$ and its standard errors, there is a trick:

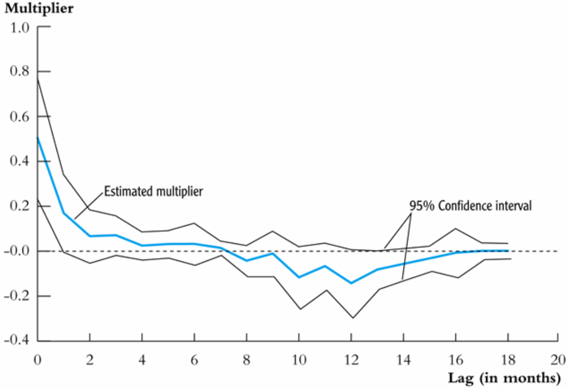
$$\begin{aligned} Y_t &= \beta_0 + \beta_1 (X_t - X_{t-1}) + [\beta_1 + \beta_2] X_{t-1} + \cdots + \beta_q X_{t-q} + u_t \\ &= \beta_0 + \beta_1 \Delta X_t + [\beta_1 + \beta_2] \Delta X_{t-1} + \cdots + [\beta_1 + \cdots + \beta_{q-1}] \Delta X_{t-} \\ &+ [\beta_1 + \cdots + \beta_q] X_{t-q} + u_t \end{aligned}$$

- where $\Delta X_t = X_t - X_{t-1}$. Simply regress Y_t on $(\Delta X_t, \ldots, \Delta X_{t-q-1})$ and X_{t-q}
- The last coefficient is the cumulative multiplier

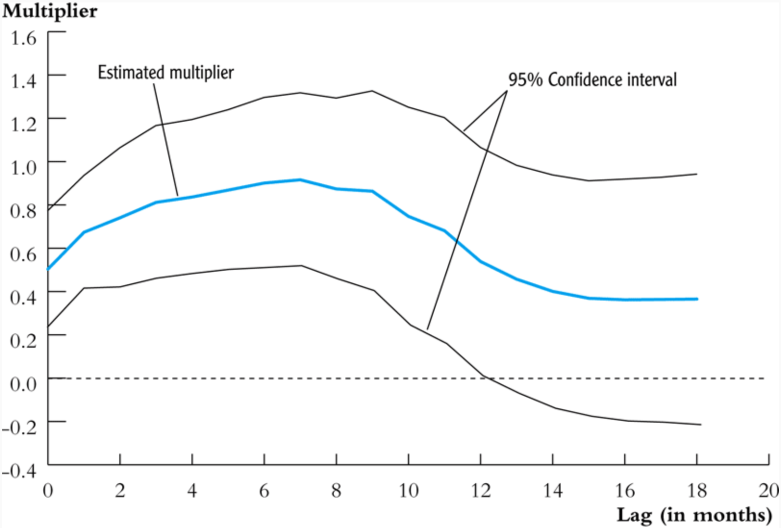
TABLE 15.1 The Dynamic Effect of a Freezing Degree Day (*FDD*) on the Price of Orange Juice: Selected Estimated Dynamic Multipliers and Cumulative Dynamic Multipliers

Lag Number	(1) Dynamic Multipliers	(2) Cumulative Multipliers	(3) Cumulative Multipliers	(4) Cumulative Multipliers
0	0.50 (0.14)	0.50 (0.14)	0.50 (0.14)	0.51 (0.15)
1	0.17 (0.09)	0.67 (0.14)	0.67 (0.13)	0.70 (0.15)
2	0.07 (0.06)	0.74 (0.17)	0.74 (0.16)	0.76 (0.18)
3	0.07 (0.04)	0.81 (0.18)	0.81 (0.18)	0.84 (0.19)
4	0.02 (0.03)	0.84 (0.19)	0.84 (0.19)	0.87 (0.20)
5	0.03 (0.03)	0.87 (0.19)	0.87 (0.19)	0.89 (0.20)
6	0.03 (0.05)	0.90 (0.20)	0.90 (0.21)	0.91 (0.21)
12 . . .	−0.14 (0.08)	0.54 (0.27)	0.54 (0.28)	0.54 (0.28)
18	0.00 (0.02)	0.37 (0.30)	0.37 (0.31)	0.37 (0.30)
Monthly indicators?	No	No	No	Yes $F = 1.01$ $(p = 0.43)$
HAC standard error truncation parameter (m)	7	7	14	7

FIGURE 15.2 The Dynamic Effect of a Freezing Degree Day (*FDD*) on the Price of Orange Juice



(a) Estimated Dynamic Multipliers and 95% Confidence Interval



(b) Estimated Cumulative Dynamic Multipliers and 95% Confidence Interval