Time Series Analysis

Intervention Analysis, Transfer Functions, and Time Series Regression

Jude C. Hays

September 11, 2024

Outline

Intervention Analysis

- Intervention models allow analysts to test for a (permanent or temporary) change in the mean of a time series at the point when an intervention occurs.
- An intervention is an event (singular or sustained) that is theorized to have an impact on the time series we are modeling.
- If we are modeling skyjackings, an intervention might be the use of metal detectors in airport security.
- If we are modeling unemployment rates, an intervention might be a switch from a right to left-wing government (or the presence of a left-wing government).
- If we are modeling President Bush's approval ratings, an intervention might be the 9/11 terrorist attacks.

Intervention Analysis

- The ARMA(p,q) version of the model is

$$y_t = a_0 + A(L)y_{t-1} + C(L)I_{t-d} + B(L)\varepsilon_t$$

where $A(L) = \left(a_1 L + a_2 L^2 + \ldots + a_p L^p\right)$,
$B(L) = \left(1 + \beta_1 L + \beta_2 L^2 + \ldots + \beta_q L^q\right)$, and
$C(L) = \left(c_d + c_{d+1} L + c_{d+2} L^2 + \ldots + c_{d+h} L^{d+h}\right)$.

- Consider a simple example. If $p = 1$, $q = 0$, $d = 0$ and $h = 0$, then we have

$$y_t = a_0 + a_1 y_{t-1} + c_0 I_t + \varepsilon_t, |a_1| < 1$$

- Without an initial condition, the solution to this model is

$$y_t = a_0/(1 - a_1) + c_0 \sum_{i=0}^{\infty} a_1^i I_{t-i} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}$$

Types of Intervention Functions

- Interventions can take several forms.

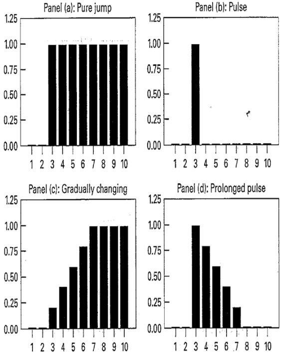


FIGURE 5.3 Typical Intervention Functions

- The step (pure jump) and pulse functions are the most commonly used.

Types of Intervention Functions

- In the case of a step function, which represents a sustained intervention, the indicator variable I_t takes a value of zero prior to the intervention and one at the time of the intervention and afterwards.
- In the case of a pulse function, which represents a singular event, the indicator variable I_t takes a value of zero in all periods except the one in which the intervention occurs.
- Ultimately, we are interested in the effect of the intervention over time. With our simple, AR(1), intervention model, the *impulse response function* for a sustained intervention is

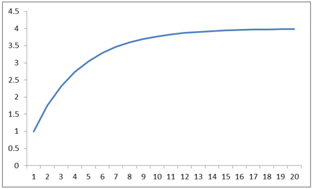
$$dy_{t+j} \Big/ dI_t = c_0 \left[1 + a_1 + ... + (a_1)^j \right]$$

The *impulse response function* for a pulse is

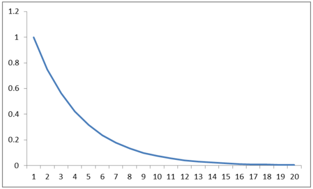
$$dy_{t+j} \Big/ dI_t = c_0 (a_1)^j$$

Impulse Response Functions

IRF: Sustained Intervention, $a_1 = .75$ & $c_0 = 1$.



IRF: Pulse Intervention, $a_1 = .75$ & $c_0 = 1$.



Estimating Intervention Models

1. Use the longest data span, either the pre or post-intervention period, to identify a set of plausible ARMA models.
2. Estimate the set of models over the entire sample including the intervention.
3. Perform diagnostics on the estimated models.

Selecting an Intervention Model

1. The coefficient estimates should be statistically significant and imply a convergent process.
2. The residuals should be white noise.
3. The model should outperform the alternatives in terms of fit and forecasting.

Transfer Function Models

- The transfer function model is a generalization of the intervention model.
- We replace the deterministic dummy variable $\{I_t\}$ with $\{z_t\}$, which can take other forms such as white-noise and ARMA processes.
- The ARMA(p,q) version of the transfer function model is

$$y_t = a_0 + A(L)y_{t-1} + C(L)z_{t-d} + B(L)\varepsilon_t$$

Identifying Transfer Function Models

- If $\{z_t\}$ is a white-noise process, we rely on the cross-correlations between $\{y_t\}$ and $\{z_{t-i}\}$ to identify the model.
- This cross-correlation is defined as

$$\rho_{yz}(i) \equiv \mathrm{cov}(y_t z_{t-i})/(\sigma_y \sigma_z)$$

.

- The standardized cross-covariance is defined as

$$\gamma_{yz}(i) \equiv \mathrm{cov}(y_t z_{t-i})/(\sigma_z^2)$$

- Plotting the cross-correlations and standardized cross-covariances give the cross-correlation function (CCF) and standardized cross-covariance function (CCVF) respectively.

Identifying Transfer Function Models

The theoretical CCF and CCVF have the following characteristics

1. All the $\rho_{yz}(i)$ and $\gamma_{yz}(i)$ will be zero until the first nonzero element of the polynomial $C(L)$.
2. The form of $B(L)$ does not affect the theoretical CCF and CCVF.
3. Spikes in the CCF and CCVF indicate a nonzero element of $C(L)$. A spike at lag d indicates that z_{t-d} directly affects y_t.
4. The nature in which spikes decay reveals information about $A(L)$. We read these patterns in the same way we read the ACF for an ARMA model.

Identifying Transfer Function Models

Some examples of theoretical CCVFs

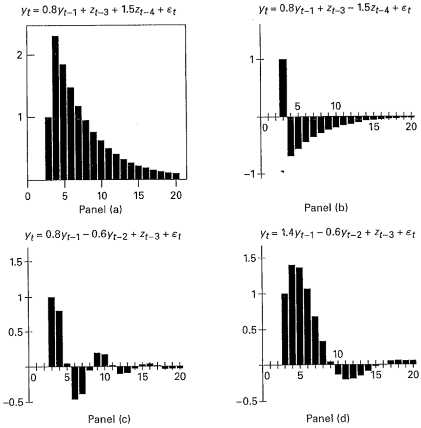


FIGURE 5.4 Standardized Cross-Correlograms

Identifying Models with Higher-Order Input Processes

- If $\{z_t\}$ is an ARMA(p,q), we have $D(L)z_t = E(L)\varepsilon_{zt}$
- Substituting into $y_t = a_0 + A(L)y_{t-1} + C(L)z_{t-d} + B(L)\varepsilon_t$ allows us to trace the effects of a one-unit shock to ε_{zt} on y_t.
- We start by estimating the parameters of the polynomials $D(L)$ and $E(L)$.
- We use the cross-correlations between $\{y_t\}$ filtered by $D(L)/E(L)$ and $\hat{\varepsilon}_{zt}$ to identify the forms of $A(L)$ and $C(L)$.
- Estimate $[1 - A(L)L]y_t = C(L)z_t + e_t$, and then use e_t to find the best model for $B(L)\varepsilon_t$
- Estimate $A(L)$, $B(L)$ and $C(L)$ simultaneously.

The Autoregressive Distributed Lag Model

We can think of time series regression as an alternative to the ARIMA approach.

- There are a number of models that we lump together under the category of time series regressions.
- The most general of these models is the autoregressive distributed lag (ADL) model

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{n} \sum_{i=0}^{q} \beta_{ji} x_{jt-i} + \varepsilon_t$$

- Consider the simple case where $p = q = n = 1$

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$$

Restricted Versions of the ADL Model

- The other common time series regression models are restricted version of the ADL.
- Among these models are

$$\text{Partial Adjustment Model: } y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \varepsilon_t$$

$$\text{Finite Distributed Lag Model: } y_t = \alpha_0 + \beta_0 x_t + \beta_1 y_{t-1} + \varepsilon_t$$

$$\text{Static Model: } y_t = \alpha_0 + \beta_0 x_t + \varepsilon_t$$

Common Factor Model:

$$y_t = \beta_0 x_t + u_t, u_t = \alpha_0 + \alpha_1 u_{t-1} + \varepsilon_t, \alpha_1 = -\beta_1/\beta_0$$

Doing Time Series Regression the Right Way

- Estimate the general model (i.e., the ADL or ECM) and test restrictions.
- Calculate all the dynamic quantities of interest (short-run effects, long-run effects, mean-lag length, median-lag length).
- Note that all of these quantities are contained in the *impulse response function*.