Regular Transfer Functions
Identification Issues and Causality Testing

Jamie Monogan
University of Georgia
January 30, 2018
Objectives

By the end of these meetings, participants should be able to:

- Explain the logic of double prewhitening and perform the procedure on real data.
- Use a cross-correlation function of double prewhitened data to identify a causal relationship.
- Estimate and diagnose a transfer function.
Where we are.

- We know how to filter series with ARIMA noise models.
- We know how to do a static regression model given the noise model.
- We know how to do a dynamic model with an intervention given the noise model.
- Now how can we specify a dynamic model with a continuous input variable given the noise model?
Dynamic and Static Again

**Dynamic** Changes in the values of $x$ affect current and future values of $y$.

**Static** Changes in the values of $x$ affect the current value of $y$.

- Introducing a lag specification for $x$ in a static model does not make it dynamic! It still has the characteristic that causation is instantaneous and complete in one time period.
- Dynamic causation flows over time.
Implied Impulse Response Function
First Order Transfer Function (Onset at Lag 1)

![Diagram showing the effect of lag on response function](image)
Implied Impulse Response Function

Koyck Scheme Regression: \( y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 x_{t-1} + e_t \)
Implied Impulse Response Function

OLS Regression: $y_t = \beta_0 + \beta_1 x_{t-1} + e_t$
The Issue

- A static regression cannot model causal flow, which is dynamic.
- It will in general produce a Type II error of inference, failing to observe an association which truly exists.
- That remains true even if you cheat and pick lag length for optimum fit.
The Effect of Introducing $x_t$ on the Right Hand Side

- A regular transfer function is exactly the model we have employed for interventions, except that $x_t$ replaces $I_t$.
- Because $x_t$ is a random variable, it will have stochastic errors.
- Because it is a time series, it is likely also to have a systematic error aggregation process.
- That presents a problem.

The standard first order case

$$y_t = \left[ \frac{\omega_0}{1 - \delta_1 B} \right] x_{t-b} + N_t \tag{1}$$

Question: How do we identify the elements of a presumed transfer function from empirical data?
A Shift of Perspective

- Terms like $a_x$ we have referred to as “errors.”
- But they are also “innovations,” and as such incorporate all of the real information in a series $x$.
- The rest of $x$ is an error aggregation process which is of no inherent interest except that we need to model it in order to control its effects.
- But in causal terms, it is the innovations, $a_x$, which matter. They are white noise by assumption—random with respect to time—but they are not random in the sense of being uncaused.
- Parallel to the information thesis of rational expectations: Only the unpredictable portion of variation represents real information in a series.
The Transfer Function Setup

Note the asymmetry in the treatment of $x$ and $y$
The Dilemma

1. \( x = f(a_x) \)
   - \( x \) is a filtered version of its innovations

2. \( y = tf(x) + N_y \)
   - our causal postulate

3. \( y = tf(f(a_x)) + N_y \)
   - substituting 1 into 2

Restated

- \( a_x \) is the causal force that drives the system.
- We want to know how \( y \) responds to \( a_x \), but we observe \( a_x \) only through its filter, \( x \).
- Solution: remove the filter.
Three Specification Decisions

- Order of $\omega$, (r)
- Order of $\delta$, (s)
- Number of periods before onset (b)

$$y_t = \left[ \frac{\omega_r}{1 - \delta_s B} \right] x_{t-b} + N_t$$  \hspace{1cm} (2)
Cross Correlation Function (CCF)

- For two variables, $x$ and $y$, the CCF is the product moment correlations between various leads and lags of the two.
- **IF** $x$ and $y$ are white noise, the CCF becomes a map of causal flow between them.
- $CCF(k) = IRF(k) \cdot (\frac{\sigma_y}{\sigma_x})$, i.e., CCF is standardized IRF.
- Given the asymmetry, $x \rightarrow y$, we need to be mindful of positive, negative, or zero lags of the CCF.
  - $CCF(0)$, is ambiguous about causality. It has no asymmetry.
  - Lags in the wrong direction could indicate the opposite causal flow, $y \rightarrow x$. 
Identifying $r$, $s$, and $b$ from the CCF

With an Example CCF

- Identify $r$, $s$, and $b$
  - $b$ first (lags before onset)
  - Then $r$—numerator order (usually 0)
    - Initial estimate of $\omega$ is $\text{IRF}(b)$
  - Then $s$—denominator order
    - Initial estimate of $\delta$ is $\text{CCF}(b+1)/\text{CCF}(b)$
Prewhitening
Single and Double Versions

**Single**
Applies the noise model for $x$, $N_x$, to both $x$ and $y$.
- Logic respects the equality of the model, applying exactly the same transformation to both sides of the equation.
- Only one side, here $x$, needs to be white noise to eliminate danger of spuriousness.
- Requires specialized software.

**Double**
Identify and estimate separate noise models for $x$ and $y$, whitening each with its own model.
- Easy to do, e.g., with R or Stata.
- Each side of the equation gets its own appropriate model.
Babies and Bathwater

- It is often said—most often by those whose cherished theories are being destroyed—that prewhitening is “throwing out the baby with the bathwater.”

- Less colorfully, that prewhitening removes some of the causal influence that we are studying and is therefore excessively conservative.

- A rebuttal: “We have so much bathwater, such as empirical findings that result from specification searches in small samples, that a baby now and then is a small cost.”

- And remember, it is only an identification technique. When we estimate a transfer function, there is no whitening applied to $x$ (except differencing, if necessary).
Exogeneity and Endogeneity Defined

- We often use the language “independent” and “dependent” to describe our variables.
- These terms describe researcher decisions—what goes on the left hand side and what goes on the right.
- But nature truly decides what is causal and what is caused.
- We need new language to describe empirical analyses of causal ordering.
- We will use *exogenous* to mean a variable that is truly uncaused in an equation system and *endogenous* to mean a variable which is caused.
- Thus causality tests are also tests for exogeneity and endogeneity.
Identification as Causality Test

- Transfer function identification thus may be seen as a causality test, telling us empirically which variables are exogenous and which endogenous.
- We are observing quite directly causal asymmetry.
- If we assert that $x \rightarrow y$, we may find evidence of that, but it may turn out empirically that
  - $y \rightarrow x$, (wrong direction) or
  - $x \Leftrightarrow y$, (no causation), or
  - $x \Leftrightarrow y$, (reciprocal causation)

Advantages and Disadvantages

- So prewhitened cross correlations are a causality test.
- Its advantage is that it shows a picture of causal asymmetry, very direct evidence.
- Its disadvantages are (1) too many steps, and (2) too much researcher judgment enters into the process.
A Preview of the Direct Granger Test

- The alternative for the bivariate case is the direct Granger test.
- In direct Granger you first regress $y$ on several lagged values of $y$, the intention being to model out all the systematic error aggregation in $y$.
- Then you add lagged values of $x$ to the model and test to see if the contribution of the $x$ lags is or is not significant.
  - Specifically, conduct a block F test.
- Then reverse the test for $y \rightarrow x$
- Direct Granger gives a clean result, $p$ values on the exogeneity of both $x$ and $y$.
- Its disadvantage, not very large, is that estimating all those extra parameters causes efficiency loss.
Step by Step Transfer Function Identification

1. Identify \( x \)
2. Estimate \( x \), and create residuals, \( \hat{e}_x \)
3. Diagnose \( x \)
   - Important! If we screw up here, all subsequent steps will be wrong.
4. Identify \( y \)
5. Estimate \( y \), and create residuals, \( \hat{e}_y \)
6. Diagnose \( y \)
   - Not as critical as \( x \) because we will have subsequent evidence.
7. Cross correlate \( x \) and \( y \) residuals
   - R: acf command (insert a matrix of two variables)
   - Stata: xcorr command
8. Identify \( r, s, \) and \( b \)
   - \( b \) first (lags before onset)
   - Then \( r \)—numerator order (usually 0). Initial estimate of \( \omega \) is IRF(\( b \))
   - Then \( s \)—denominator order. Initial estimate of \( \delta \) is CCF(\( b+1 \))/CCF(\( b \))
9. Estimate transfer function
10. Diagnosis: to be discussed
Critiques of Box-Jenkins Models

- Too much art:
  - Two competent analysts will often produce different models.

- Atheoretical:
  - Often just wrong, confusing structure (TF) with error (ARIMA).
  - Sometimes cheap shot, atheoretical in situations where there is no theory.

- Identification is fundamentally bivariate:
  - There is no multivariate procedure.

- The baby and the bathwater:
  - Might prewhitening eliminate evidence of genuine causality (along with spuriousness)?
An Important Example

- Long and short rates defined
- We have a clear theory. Long rates are determined by short rates and by inflation expectations.
  - Long = f(Short, Expectations)
- So we should be able to identify and estimate the impact of short rates on long rates.
Problems with Prewhitening (and Granger Tests)

Long and Short Rates

![Graph showing Long and Short Interest Rates](image)

- **X-axis:** Year, from Jan-46 to Jan-07
- **Y-axis:** Percent
- **Legend:**
  - Black line: Treasury Bonds
  - Pink line: Prime Rate
ARIMA modeling assumes that variance is a constant $\sigma^2$ for all $t$ (homoscedasticity).

Some series, such as these, have the property that variance is related to the level of the series.

- As the level changes, e.g., growing with price inflation, the variance increases accordingly.

Heteroscedasticity, $\sigma_i^2 \neq \sigma_j^2$ for $i \neq j$, produces estimator inefficiency.

If constant percentage growth is the problem, then logging the series is the fix.

(Preview: if logging is not sufficient, then ARCH/GARCH.)
The Two Series as Logs
Problems with Prewhitening (and Granger Tests)

Cross Correlation Function. This Figure is Opposite of R. Short Rates ($x=\text{Prime}$) leading Long Rates ($y=\text{Tbills}$)
Conclusion

- Short rates do not cause long rates. Therefore
  - Our theory is wrong, and
  - The covariation we think we see between them must be false.
- But this can’t be! Short rates must cause long rates.
- The baby and the bathwater again.
  - The method must be wrong.
- To be addressed later in the course.
Transfer Function Diagnosis

- **Two Questions:**
  - (1) Do we have the right noise model for y?
  - (2) Have we captured the flow of x into y?
- (1) The usual residual white noise test
  - If failed, then respecify noise model for y
- (2) Cross correlation of prewhitened x with TF residuals
  - If there is any significant correlation, the TF specification is wrong.
- Break to R to work through a transfer function.
For Next Time:

- Download Enders’s (2009) monetary policy data `moneyDem.dta` from http://dx.doi.org/10.7910/DVN/ARKOTI.
- A description of these data can be found on page 186 as part of question #4 in *Political Analysis Using R*. Do not attempt to work question #4. Just feel free to consult the variable information.
- Consider change in the real M2 money supply (`d1rm2`) and the 3-month interest rate on U.S. Treasury bills (`drs`).
- Complete all of the steps of identifying, estimating, and diagnosing a transfer function between these variables. Prewhiten each. Choose a specification. Check that it is valid.
- Warning: One variable’s ARIMA process is hard to nail. Trying a few things I still got a similar transfer function result.

**Reading:**