Estimation of Dynamic Causal Effects

Ming-Ching Luoh

Dynamic Causal Effects

Estimation with Exogenous Regressors

HAC Standard Errors

Estimation with Strictly Exogenous Regressors

Orange Juice Prices and Cold Weather

Is Exogeneity Plausible?

Dynamic Causal Effect

A **dynamic causal effect** is the effect on Y of a change in X over time.

For example,

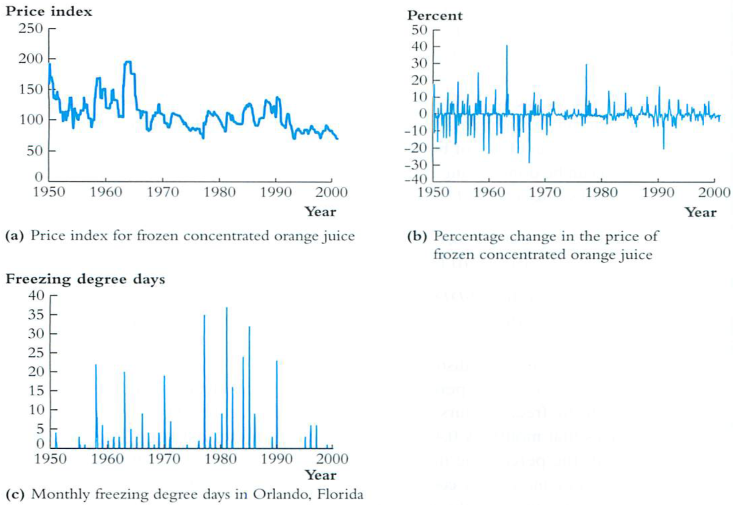
- The effect of an increase in cigarette taxes on cigarette consumption this year, next year, in 5 years.
- The effect of a change in the Fed Funds rate on inflation, this month, in 6 months, and 1 year.
- The effect of a freeze in Florida on the price of orange juice concentrate in 1 month, 2 months, 3 months $\cdots$.

The Orange Juice Data

- Monthly, Jan. 1950 - Dec. 2000 (T = 612).
- Price = price of frozen OJ concentrate (a sub-component of the producer price index; US Bureau of Labor Statistics)
- $\%ChgP$ = percentage change in price at an annual rate, so $\%ChgP_t = 12 \cdot 100 \cdot \Delta \ln(Price_t)$.
- FDD = sum of the number of degrees in the "freezing degree-days" that the minimum temperatures falls below 32^o during the month, recorded in Orlando, FL.
    - Example: If November has 2 days with low temp $< 32^o$, one at 30^o and at 25^o, then

$$FDD_{Nov} = 2 + 7 = 9.$$

FIGURE 16.1 Orange Juice Prices and Florida Weather, 1950–2000



There have been large month-to-month changes in the price of frozen concentrated orange juice. Many of the large movements coincide with freezing weather in Orlando, home of many orange groves.

Initial OJ regression

$$\%ChgP_t = \underset{(.22)}{-.40} + \underset{(.13)}{.47}\, FDD_t$$

- Statistically significant positive relation.
- More/deeper freezes, price goes up.
- Standard errors: not the usual OLS s.e., but rather are heteroskedasticity- and autocorrelation-consistent $(HAC) SE$'s - *more on this later*.
- But what is the effect of FDD over time?

Dynamic Causal Effects

Example: What is the effect of fertilizer on tomato yield?

An ideal randomized controlled experiment.

- Fertilize some plots, not others (random assignment).
- Measure yield *over time - over repeated harvests* - to estimate causal effect of fertilizer on
  - Yield in year 1 of experiment
  - Yield in year 2, etc.
- The result is the causal effect of fertilizer on yield k years later.

In time series applications, we can't conduct this ideal randomized controlled experiment:

- We only have one U.S. OJ market.
- We can't randomly assign FDD to different replicates of the U.S. OJ market.
- So we can't estimate the causal effect at different times using the differences estimator.

An alternative thought experiment:

- Randomly give the same subject different treatments (FDD_t) *at different times*.
- Measure the outcome variable ($\%ChgP_t$)
- The "population" of subjects consists of the same subject (OJ market) but at different dates.
- If the "different subjects" are drawn from the same distribution - that is, if Y_t, X_t are stationary - then the dynamic causal effect can be deduced by OLS regression of Y_t on lagged values of X_t.
- This estimator (regression of Y_t on X_t and lags of X_t's) called the *distributed lag* estimator.

Dynamic causal effects and the distributed lag model

The **distributed lag model** is:

$$Y_t = \beta_0 + \beta_1 X_t + \cdots + \beta_{r+1} X_{t-r} + u_t$$

- β_1 = impact effect of change in X = effect of change in X_t on Y_t, holding past X_t constant.
- β_2 = 1-period dynamic multiplier = effect of change in X_{t-1} on Y_t, holding constant $X_t, X_{t-2}, X_{t-3}, \cdots$.
- β_3 = 2-period dynamic multiplier (etc.)= effect of change in X_{t-2} on Y_t, holding constant $X_t, X_{t-1}, X_{t-3}, \cdots$.
- *Cumulative dynamic multipliers*
  - Ex: the 2-period cumulative dynamic multiplier = $\beta_1 + \beta_2 + \beta_3$

Exogeneity in time series regression

- **Exogeneity** (past and present)
  X is exogenous if $\mathrm{E}(u_t|X_t, X_{t-1}, X_{t-2}, \cdots) = 0$.
- **Strict Exogeneity** (past, present, and future) X is strictly exogenous if $\mathrm{E}(u_t|\cdots, X_{t+1}, X_t, X_{t-1}, \cdots) = 0$.
- Strict exogeneity implies exogeneity.
- For now we suppose that X is exogenous - we'll return (briefly) to the case of strict exogeneity later.
- If X is exogenous, then OLS estimates the dynamic causal effect on Y of a change in X.

Estimation of Dynamic Causal Effects with Exogenous Regressors

$$Y_t = \beta_0 + \beta_1 X_t + \cdots + \beta_{r+1} X_{t-r} + u_t$$

The Distributed Lag Model Assumptions

1. $\mathrm{E}(u_t | X_t, X_{t-1}, X_{t-2}, \cdots) = 0$ (X is exogenous).
2. (a) Y and X have stationary distributions.
   (b) (Y_t, X_t) and (Y_{t-j}, X_{t-j}) become independent as j gets large.
3. Y and X have eight nonzero finite moments.
4. There is no perfect multicollinearity.

- Assumptions 1 and 4 are familiar.
- Assumption 3 is familiar, except for 8 (not four) finite moments - this has to do with HAC estimators.
- Assumption 2 is different - before it was (X_i, Y_i) are i.i.d. - this now becomes more complicated.

Assumption 2(a): Y and X have stationary distributions.

- If so, the coefficients don't change within the sample (internal validity),
- and the results can be extrapolated outside the sample (external validity).
- This is the time series counterpart of the "identically distributed" part of i.i.d.

Assumption 2(b): (Y_t, X_t) and (Y_{t-j}, X_{t-j}) become independent as j gets large.

- Intuitively, this says that we have separate experiments for time periods that are widely separated.
- In cross-sectional data, we assumed that Y and X were i.i.d., a consequence of simple random sampling - this led to the CLT.
- A version of the CLT holds for time series variables that become independent as their temporal separation increases - assumption 2(b) is the time series counterpart of the "independently distributed" part of i.i.d.

Under the Distributed Lag Model Assumptions:

- OLS yields consistent estimators of $\beta_1, \beta_2, \cdots, \beta_r$ (of the dynamic multipliers).
- The sampling distribution of $\hat{\beta}_1$, etc., is normal.
- *However*, the formula for the variance of this sampling distribution is not the usual one from cross-sectional (i.i.d.) data, because u_t is not i.i.d. - it is serially correlated.
- This means that the usual OLS standard errors (usual STATA printout) are wrong.
- We need to use, instead, SEs that are robust to autocorrelation as well as to heteroskedasticity.

Heteroskedasticity and Autocorrelation-Consistent (HAC) Standard Errors

- When u_t is serially correlated, the variance of the sampling distribution of the OLS depends on this serial correlation.
- Consequently, we need to use a different formula for the standard errors.
- This is easy to do using STATA and most other statistical software.

The math

Consider the case of no lags:

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

Recall that the OLS estimator is:

$$\hat{\beta}_1 = \frac{\frac{1}{T}\sum_{t=1}^{T}(X_t - \bar{X})(Y_t - \bar{Y})}{\frac{1}{T}\sum_{t=1}^{T}(X_t - \bar{X})^2}$$

so, in large samples,

$$\hat{\beta}_1 - \beta_1 = \frac{\frac{1}{T}\sum_{t=1}^{T}(X_t - \bar{X})u_t}{\frac{1}{T}\sum_{t=1}^{T}(X_t - \bar{X})^2} \cong \frac{\frac{1}{T}\sum_{t=1}^{T} v_t}{\sigma_X^2}$$

where $v_t = (X_t - \bar{X})u_t$.

Thus, in large samples,

$$\begin{aligned}\mathrm{Var}(\hat{\beta}_1) &= \frac{\mathrm{Var}(\frac{1}{T}\sum_{t=1}^{T} v_t)}{(\sigma_X^2)^2} \\ &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{\mathrm{Cov}(v_t, v_s)}{(\sigma_X^2)^2}\end{aligned}$$

In i.i.d. cross sectional data, $\mathrm{Cov}(v_t, v_s) = 0$ for $t \neq s$, so

$$\mathrm{Var}(\hat{\beta}_1) = \frac{1}{T^2}\frac{T\mathrm{Var}(v_t)}{(\sigma_X^2)^2} = \frac{\sigma_v^2}{T(\sigma_X^2)^2}$$

This is the usual cross-sectional result.

What happens with time series data?

Consider T = 2:

$$\begin{aligned}
\text{Var}\left(\frac{1}{T}\sum_{t=1}^{T} v_t\right) &= \text{Var}\left(\frac{1}{2}(v_1+v_2)\right) \\
&= \frac{1}{4}\left(\text{Var}(v_1) + \text{Var}(v_2) + 2\text{Cov}(v_1, v_2)\right) \\
&= \frac{1}{2}\sigma_v^2 + \frac{1}{2}\rho_1\sigma_v^2, \quad (\rho_1 = corr(v_1, v_2)) \\
&= \frac{1}{2}\sigma_v^2 \times f_2, \text{ where } f_2 \equiv (1+\rho_1)
\end{aligned}$$

- In i.i.d. (cross-section) data, $\rho_1 = 0$ and $f_2 = 1$, which gives the usual formula.
- In time series data, $\rho_1 \neq 0$, so $\text{Var}(\hat{\beta}_1)$ is not given by the usual formula.

Expression for $\text{Var}(\hat{\beta}_1)$, general T

$$\text{Var}\left(\frac{1}{T}\sum_{t=1}^{T} v_t\right) = \frac{\sigma_v^2}{T} \times f_T$$

$$\text{Var}(\hat{\beta}_1) = \left(\frac{1}{T}\frac{\sigma_v^2}{(\sigma_X^2)^2}\right) \times f_T$$

where

$$f_T = 1 + 2\sum_{j=1}^{T-1}\left(\frac{T-j}{T}\right)\rho_j$$

- Conventional OLS *SE*'s are wrong when u_t is serially correlated (STATA printout is wrong).
- The OLS *SE*'s are off by the factor f_T (which can be big!)
- We need to use a different *SE* formula!!!

HAC Standard Errors

- Conventional OLS SEs (heteroskedasticity-robust or not) are wrong when there is autocorrelation.
- So, we need a new formula that produces SEs that are robust to autocorrelation as well as heteroskedasticity. **We need Heteroskedasticity and Autocorrelation-Consistent (HAC) standard errors**.
- If we knew the factor f_T, we could just make the adjustment.
- But we don't know f_T - it depends on unknown autocorrelations, ρ_1, ρ_2.
- HAC SEs replace f_T with an estimator of f_T.

$$\text{Var}(\hat{\beta}_1) = \left(\frac{1}{T} \frac{\sigma_v^2}{(\sigma_X^2)^2} \right) \times f_T,$$

where

$$f_T = 1 + 2 \sum_{j=1}^{T-1} \left(\frac{T-j}{T} \right) \rho_j$$

The most commonly used estimator of f_T is

$$\hat{f}_T = 1 + 2 \sum_{j=1}^{m-1} \left(\frac{m-j}{m} \right) \tilde{\rho}_j$$

$$\hat{f}_T = 1 + 2\sum_{j=1}^{m-1}\left(\frac{m-j}{m}\right)\tilde{\rho}_j$$

- This is sometimes called "Newey-West" weights.
- $\tilde{\rho}_j$ is an estimator of ρ_j.
- m is called the **truncation parameter**.
- What truncation parameter to use in practice such that $\hat{f}_T$ is consistent?
    - Try $m = 0.75T^{1/3}$, rounded to an integer.

Example: OJ and HAC estimators in STATA

```
. gen l0fdd = fdd;                  generate lag #0
. gen l1fdd = L1.fdd;               generate lag #1
. gen l2fdd = L2.fdd;               generate lag #2
. gen l3fdd = L3.fdd;                   .
. gen l4fdd = L4.fdd;                   .
. gen l5fdd = L5.fdd;                   .
. gen l6fdd = L6.fdd;

. reg dlpoj fdd if tin(1950m1,2000m12), r;      NOT HAC SEs

Linear regression                                      Number of obs =     612
                                                       F(  1,   610) =   12.12
                                                       Prob > F      =  0.0005
                                                       R-squared     =  0.0937
                                                       Root MSE      =  4.8261

------------------------------------------------------------------------------
             |               Robust
       dlpoj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fdd |   .4662182   .1339293     3.48   0.001     .2031998    .7292367
       _cons |  -.4022562   .1893712    -2.12   0.034    -.7741549   -.0303575
------------------------------------------------------------------------------
```

Example: OJ and HAC estimators in STATA, ctd

Rerun this regression, but with Newey-West *SEs*:

```
.  newey dlpoj fdd if tin(1950m1,2000m12), lag(7);

Regression with Newey-West standard errors          Number of obs  =       612
maximum lag: 7                                      F(  1,   610)  =     12.23
                                                    Prob > F       =    0.0005

------------------------------------------------------------------------------
             |             Newey-West
       dlpoj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fdd |   .4662182   .1333142     3.50   0.001     .2044077    .7280288
       _cons |  -.4022562   .2159802    -1.86   0.063    -.8264112    .0218987
------------------------------------------------------------------------------
```

Uses autocorrelations up to m = 7 to compute the SEs

rule-of-thumb: $0.75*(612^{1/3}) = 6.4$ □ 7, rounded up a little.

OK, in this case the difference in SEs is small, but not always so!

Example: OJ and HAC estimators in STATA, ctd.

```
. global lfdd6 "fdd l1fdd l2fdd l3fdd l4fdd l5fdd l6fdd";

.  newey dlpoj $lfdd6 if tin(1950m1,2000m12), lag(7);

Regression with Newey-West standard errors          Number of obs  =       612
maximum lag : 7                                     F(  7,   604)  =      3.56
                                                    Prob > F       =    0.0009

------------------------------------------------------------------------------
             |             Newey-West
       dlpoj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fdd |   .4693121   .1359686     3.45   0.001     .2022834    .7363407
       l1fdd |   .1430512   .0837047     1.71   0.088    -.0213364    .3074388
       l2fdd |   .0564234   .0561724     1.00   0.316    -.0538936    .1667404
       l3fdd |   .0722595   .0468776     1.54   0.124    -.0198033    .1643223
       l4fdd |   .0343244   .0295141     1.16   0.245    -.0236383    .0922871
       l5fdd |   .0468222   .0308791     1.52   0.130    -.0138212    .1074657
       l6fdd |   .0481115   .0446404     1.08   0.282    -.0395577    .1357807
       _cons |  -.6505183   .2336986    -2.78   0.006    -1.109479   -.1915578
------------------------------------------------------------------------------
```

- global lfdd6 *defines a string which is all the additional lags*
- *What are the estimated dynamic multipliers (dynamic effects)?*

Do I need to use HAC SEs when I estimate an AR or an ADL model?

NO.

- The problem to which HAC SEs are the solution arises when u_t is serially correlated.
- If u_t is serially uncorrelated, then OLS SE's are fine.
- In AR and ADL models, the errors are serially uncorrelated if you have included enough lags of Y.
  - If you include enough lags of Y, then the error term can't be predicted using past Y, or equivalently by past u - so u is serially uncorrelated.

Estimation of Dynamic Causal Effects with Strictly Exogenous Regressors

- X is strictly exogenous if $E(u_t | \cdots, X_{t+1}, X_t, X_{t-1}, \cdots) = 0$.
- If X is strictly exogenous, there are more efficient ways to estimate dynamic causal effects than by a distributed lag regression.
  - Generalized Least Squares (GLS)
  - Autoregressive Distributed Lag (ADL)

- But the condition of strict exogeneity is very strong, so this condition is rarely plausible in practice.
- So we won't cover GLS or ADL estimation of dynamic causal effects— for details, see SW Section 15.5.

Orange Juice Prices and Cold Weather

What is the dynamic causal effect (*what are the dynamic multipliers*) of a unit increase in FDD on OJ prices?

$$\%ChgP_t = \beta_0 + \beta_1 FDD_t + \cdots + \beta_{r+1} FDD_{t-r} + u_t$$

- What r to use? How about 18?
- What m (Newey-West truncation parameter) to use? $m = .75 \times 612^{1/3} = 6.4 \cong 7$.

Digression: Computation of cumulative multipliers and their standard errors

- The cumulative multipliers can be computed by estimating the distributed lag model, then adding up the coefficients. However, you should also compute standard errors for the cumulative multipliers and while this can be done directly from the distributed lag model it requires some modifications.
- Because cumulative multipliers are linear combinations of regression coefficients, the methods of Section 7.3 can be applied to compute their standard errors.
- A trick in Section 7.3 is to rewrite the regression so that the coefficients in the rewritten regression are the coefficients of interest— here, the cumulative multipliers.

Rewrite the distributed lag model with 1 lag:

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + u_t \\ &= \beta_0 + \beta_1 X_t - \beta_1 X_{t-1} + \beta_1 X_{t-1} + \beta_2 X_{t-1} + u_t \\ &= \beta_0 + \beta_1 (X_t - X_{t-1}) + (\beta_1 + \beta_2) X_{t-1} + u_t \\ Y_t &= \beta_0 + \beta_1 \Delta X_t + (\beta_1 + \beta_2) X_{t-1} + u_t \end{aligned}$$

So, let $W_{1t} = \Delta X_t$ and $W_{2t} = X_{t-1}$ and estimate the regression,

$$Y_t = \beta_0 + \delta_1 W_{1t} + \delta_2 W_{2t} + u_i$$

Then

$$\begin{aligned} \delta_1 &= \beta_1 = impact\, effect \\ \delta_2 &= \beta_1 + \beta_2 = the\, first\, cumulative\, multiplier \end{aligned}$$

and the (HAC) standard errors on δ_1 and δ_2 are the standard errors for the two cumulative multipliers.

In general, the ADL model can be rewritten as,

$$Y_t = \gamma_0 + \gamma_1 \Delta X_t + \gamma_2 \Delta X_{t-1} + \cdots + \gamma_{q-1} \Delta X_{t-q+1} + \gamma_q X_{t-q} + u_t$$

where $\gamma_1 = \beta_1$, $\gamma_2 = \beta_1 + \beta_2$, $\gamma_3 = \beta_1 + \beta_2 + \beta_3$, etc.

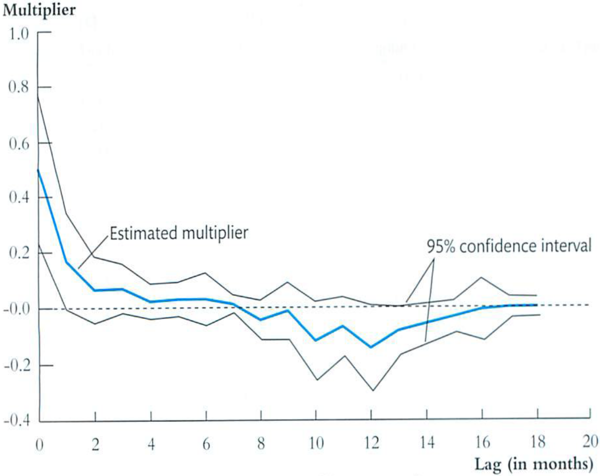
- Cumulative multipliers and their HAC SEs can be computed directly using this transformed regression.

TABLE 16.1 The Dynamic Effect of a Freezing Degree Day (*FDD*) on the Price of Orange Juice: Selected Estimated Dynamic Multipliers and Cumulative Dynamic Multipliers

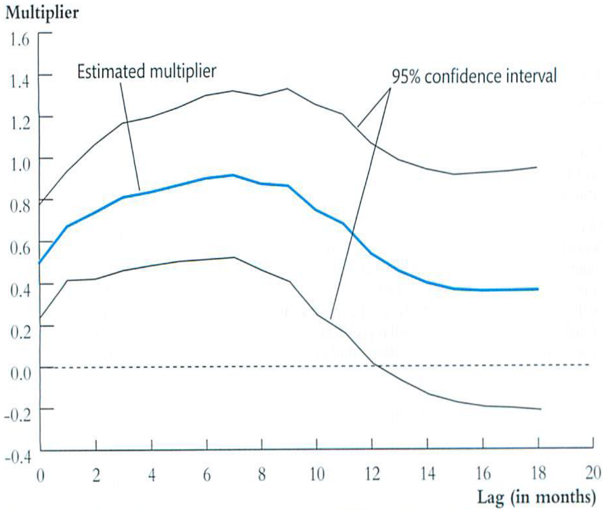
Lag Number	(1) Dynamic Multipliers	(2) Cumulative Multipliers	(3) Cumulative Multipliers	(4) Cumulative Multipliers
0	0.50 (0.14)	0.50 (0.14)	0.50 (0.14)	0.51 (0.15)
1	0.17 (0.09)	0.67 (0.14)	0.67 (0.13)	0.70 (0.15)
2	0.07 (0.06)	0.74 (0.17)	0.74 (0.16)	0.76 (0.18)
3	0.07 (0.04)	0.81 (0.18)	0.81 (0.18)	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 16.2 The Dynamic Effect of a Freezing Degree Day (*FDD*) on the Price of Orange Juice

The estimated dynamic multipliers show that a freeze leads to an immediate increase in prices. Subsequent price rises are much smaller than the initial impact. The cumulative multiplier shows that freezes have a persistent effect on the level of orange juice prices, with prices peaking seven months after the freeze.



(a) Estimated dynamic multipliers and 95% confidence interval



(b) Estimated cumulative dynamic multipliers and 95% confidence interval

Are the OJ dynamic effects stable?

Compute QLR for regression (1) in Table 15.1:

- Do you need HAC SEs?
- How specifically would you compute the Chow statistic?
- How would you compute the QLR statistic?
- What are the d.f. q of the QLR statistics?
  $q = 20$, FDD_t, 18 lags and intercept.
- Result QLR = 21.19.
  - Is this significant?
  - At what significance level?
  - The 1% critical value in Table 14.6 is 2.43.
- Estimate the dynamic multipliers on subsamples and see how they have changed over time.

OJ: Do the breaks matter substantively?

TABLE 15.5 Critical Values of the QLR Statistic with 15% Trimming

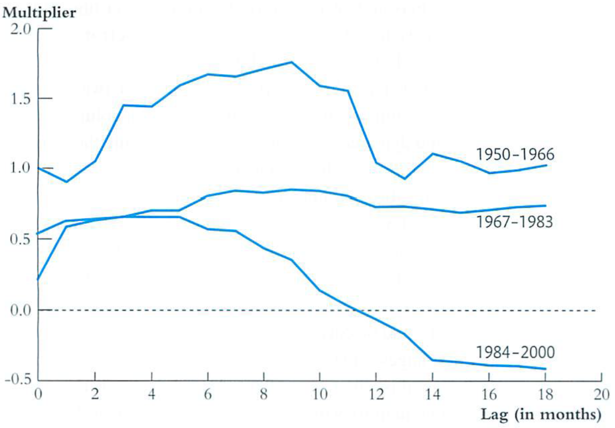
Number of Restrictions (q)	10%	5%	1%
1	7.12	8.68	12.16
2	5.00	5.86	7.78
3	4.09	4.71	6.02
4	3.59	4.09	5.12
5	3.26	3.66	4.53
6	3.02	3.37	4.12
7	2.84	3.15	3.82
8	2.69	2.98	3.57
9	2.58	2.84	3.38
10	2.48	2.71	3.23
11	2.40	2.62	3.09
12	2.33	2.54	2.97
13	2.27	2.46	2.87
14	2.21	2.40	2.78
15	2.16	2.34	2.71
16	2.12	2.29	2.64
17	2.08	2.25	2.58
18	2.05	2.20	2.53
19	2.01	2.17	2.48
20	1.99	2.13	2.43

Note: These critical values apply when $\tau_0 = 0.15T$ and $\tau_1 = 0.85T$ (rounded to the nearest integer), so the F-statistic is computed for all potential break dates in the central 70% of the sample. The number of restrictions q is the number of restrictions tested by each individual F-statistic. Critical values for other trimming percentages are given in Andrews (2003).

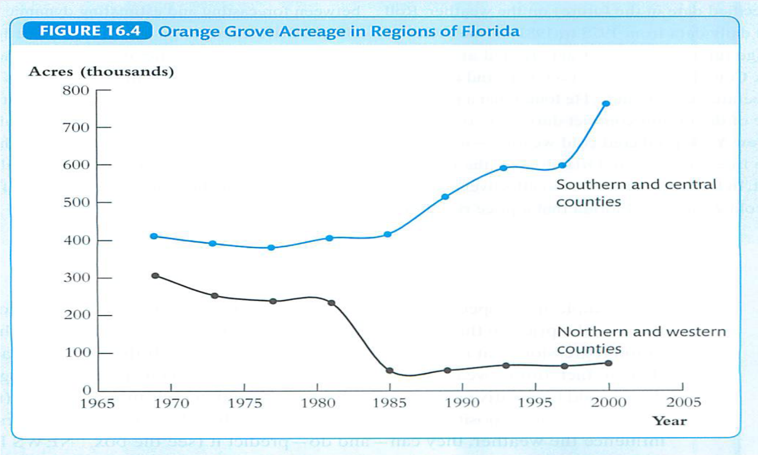
OJ: Do the breaks matter substantively?

FIGURE 16.3 Estimated Cumulative Dynamic Multipliers from Different Sample Periods

The dynamic effect on orange juice prices of freezes changed significantly over the second half of the 20th century. A freeze had a larger impact on prices during 1950–1966 than later, and the effect of a freeze was less persistent during 1984–2000 than earlier.



- The cumulative effect of a FDD declines over time? Why?



- After losing many trees to freezes in northern Florida, Florida orange growers planted farther south. *How does this relate to the change in the cumulative impulse responses*?

These dynamic multipliers were estimated using a distributed lag model. Should we attempt to obtain more efficient estimates using GLS or an ADL model?

- Is FDD strictly exogenous in the distributed lag regression?

$$\%ChgP_t = \beta_0 + \beta_1 FDD_t + \cdots + \beta_{r+1} FDD_{t-r} + u_t$$

- OJ commodity traders can't change the weather.
- So this implies that $\text{Corr}(u_t, FDD_{t+1}) = 0$, right?
- No. Although traders can not change the weather, they can predict it. Therefore, price today will be affected by weather in the future. $\text{Corr}(u_t, FDD_{t+1}) \neq 0$.

Is Exogeneity Plausible?

When can you estimate dynamic causal effects? That is, when is exogeneity plausible?

In the following examples,

- is X exogenous?
- is X strictly exogenous?

Examples:

1. Y = OJ prices, X = FDD in Orlando.
2. Y = Australian exports, X = US GDP (effect of US income on demand for Australian exports)

Examples, ctd.

3. Y = EU exports, X = US GDP (effect of US income on demand for EU exports)

4. Y = US rate of inflation, X = percentage change in world oil prices (as set by OPEC) (effect of OPEC oil price increase on inflation)

5. Y = GDP growth, X =Federal Funds rate (the effect of monetary policy on output growth)

6. Y = change in the rate of inflation, X = unemployment rate on inflation (the Phillips curve)

Exogeneity, ctd.

- You must evaluate exogeneity and strict exogeneity on a case by case basis.
- Exogeneity is often not plausible in time series data because of simultaneous causality.
- Strict exogeneity is rarely plausible in time series data because of feedback.

Estimation of Dynamic Causal Effects: Summary

- Dynamic causal effects are measurable in theory using a randomized controlled experiment with repeated measurements over time.
- When X is exogenous, you can estimate dynamic causal effects using a distributed lag regression.
- If u_t is serially correlated, conventional OLS SEs are incorrect. You must use HAC SEs.
- To decide whether X is exogenous, think hard!