

# Estimación de Efectos Causales Dinámicos

## (SW Capítulo 13)

Un *efecto causal dinámico* es el que produce un cambio en  $X$  sobre  $Y$  a lo largo del tiempo.

Por ejemplo:

- El efecto de un incremento en los impuestos del tabaco sobre su consumo de este año, del próximo, o del de dentro de cinco años.
- El efecto de un cambio en el tipo de interés sobre la inflación de este mes, de 6 meses o de un año.
- El efecto de una helada en Valencia sobre el precio del zumo de naranja en 1 mes, 2 meses, 3 meses...

# Datos de Zumo de Naranja

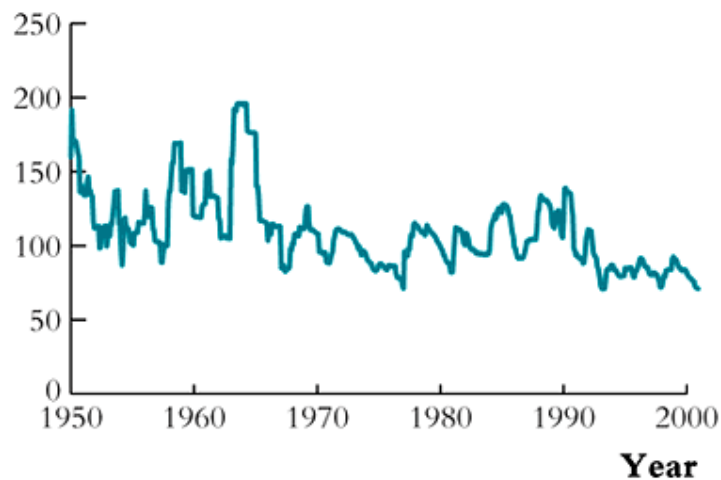
## (SW Sección 13.1)

### Datos

- Mensuales, Enero 1950 – Diciembre 2000 ( $T = 612$ )
- $Precio$  = precio del zumo
- $\%ChgP$  = tasa anual de cambio porcentual en el precio; es decir,  $ChgP_t = 1200\Delta\ln(Precio_t)$
- $FDD$  = número de grados-día bajo cero en un mes
  - Ejemplo: Si en noviembre hay 2 días bajo cero, uno a  $-2^\circ$  y otro a  $-3^\circ$ , entonces  $FDD_{Nov} = 2 + 3 = 5$

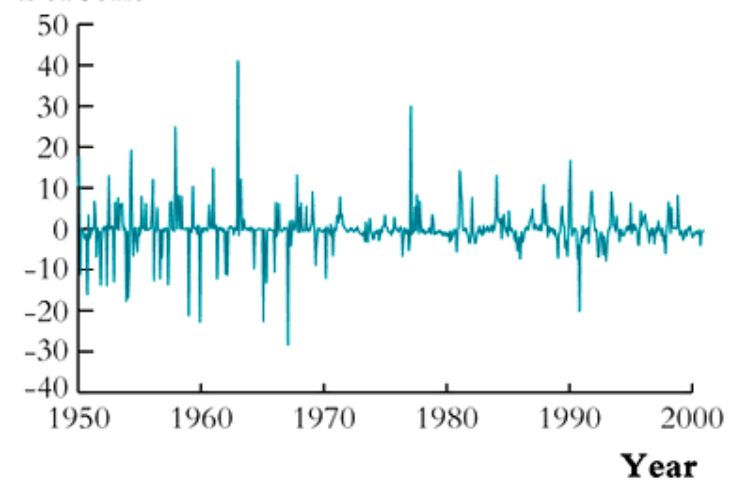
**FIGURE 13.1** Orange Juice Prices and Florida Weather, 1950–2000

**Price Index**



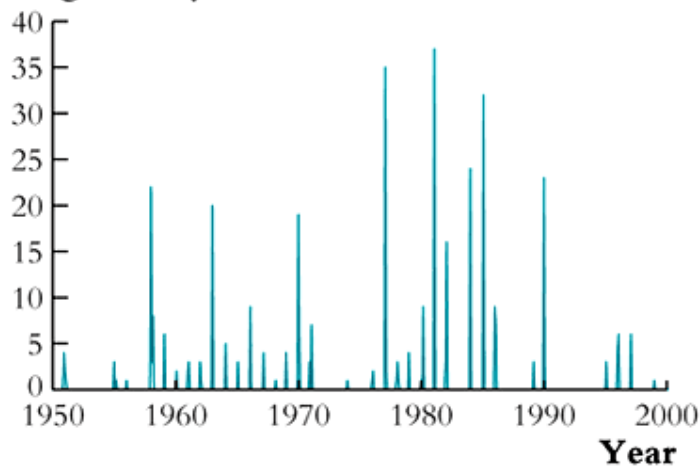
**(a)** Price Index for Frozen Concentrated Orange Juice

**Percent**



**(b)** Percent Change in the Price of Frozen Concentrated Orange Juice

**Freezing Degree Days**



**(c)** Monthly Freezing Degree Days in Orlando, Florida

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 the orange groves.

## Regresión inicial

$$\widehat{\%ChgP}_t = -.40 + .47FDD_t$$

(.22) (.13)

- Relación positiva estadísticamente significativa.
- A más frío, mayor precio
- Errores estándar *consistentes a la heteroscedasticidad y autocorrelación (HAC)*
- ¿Cuál es el efecto de *FDD* a lo largo del tiempo?

# Efectos Causales Dinámicos

## (SW Sección 13.2)

*Ejemplo:* ¿Cuál es el efecto de un fertilizante sobre el rendimiento en la producción de tomates?

Experimento aleatorizado controlado

- Fertilizar sólo en algunas partes (asignación aleatoria)
- Medir el rendimiento de las cosechas a lo largo del tiempo – estimar el efecto causal del fertilizante en
  - El rendimiento de 1 año
  - El rendimiento dentro de 2 años, etc.

En aplicaciones de series temporales, no podemos llevar a cabo experimentos en condiciones ideales

- Sólo disponemos de un único mercado.
- No podemos asignar aleatoriamente FDD a diferentes replicas del mercado.
- No podemos medir el resultado *medio* en diferentes momentos del tiempo.
- Por tanto, no podemos estimar el efecto causal en diferentes momentos del tiempo mediante el estimador de diferencias.

## Experimento alternativo:

- Aleatoriamente dar al mismo sujeto (mercado de zumos) distintos tratamientos ( $FDD_t$ ) en diferentes momentos del tiempo.
- Medir el resultado ( $\%ChgP_t$ )
- La “población” de sujetos consiste en el mismo sujeto en diferentes momentos del tiempo
- Si los “distintos sujetos” obedecen la misma ley – *si  $Y_t, X_t$  son estacionarias* – entonces el efecto causal dinámico puede obtenerse a través de la regresión MCO de  $Y_t$  sobre valores retardados de  $X_t$ .
- Este estimador se denomina de *retardos distribuidos*.

## Modelo de retardos distribuidos

$$Y_t = \beta_0 + \beta_1 X_t + \dots + \beta_p X_{t-r} + u_t$$

- $\beta_1 = \textit{impacto de un cambio en } X = \text{efecto de un cambio en } X_t \text{ sobre } Y_t, \text{ manteniendo constante el pasado de } X_t$
- $\beta_2 = \textit{multiplicador dinámico de 1-periodo} = \text{efecto de un cambio en } X_{t-1} \text{ sobre } Y_t, \text{ manteniendo constante } X_t, X_{t-2}, X_{t-3}, \dots$
- $\beta_3 = \textit{multiplicador dinámico de 2-periodos} = \text{efecto de un cambio en } X_{t-2} \text{ sobre } Y_t, \text{ manteniendo constante } X_t, X_{t-1}, X_{t-3}, \dots$
- ***Multiplicadores dinámicos acumulados***
  - *Ej.:* el multiplicador dinámico acumulado de 2-periodos es  $\beta_1 + \beta_2 + \beta_3$



# Exogeneidad en regresiones de series temporales

## Exogeneidad (*pasada y presente*)

$X$  es *exógena* si  $E(u_t | X_t, X_{t-1}, X_{t-2}, \dots) = 0$ .

## Exogeneidad estricta (*pasada, presente, y futura*)

$X$  es *estrictamente exógena* si

$$E(u_t | \dots, X_{t+1}, X_t, X_{t-1}, \dots) = 0$$

- Exogeneidad estricta implica exogeneidad
- Supongamos por ahora que  $X$  es exógena
- Si  $X$  es exógena entonces MCO estima el efecto causal dinámico de un cambio en  $X$  sobre  $Y$

# Estimación de Efectos causales dinámicos con regresores exógenos (SW Sección 13.3)

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

## Supuestos del modelo de retardos distribuidos

1.  $E(u_t | X_t, X_{t-1}, X_{t-2}, \dots) = 0$  ( $X$  es exógena)
2. (a)  $Y, X$  son estacionarias;  
(b)  $(Y_t, X_t)$  and  $(Y_{t-j}, X_{t-j})$  se convierten en independientes cuando  $j$  tiende a infinito.
3.  $Y, X$  poseen 8 momentos finitos no nulos.
4. No se da multicolinealidad perfecta.

- Los supuestos 1 y 4 son ya conocidos
- El supuesto 3 es familiar, excepto por el hecho de ser 8 (y no 4) los momentos finitos necesarios (tiene que ver con el estimador HAC)
- El supuesto 2 es diferente – antes  $(X_i, Y_i)$  eran i.i.d.

(a)  $Y, X$  son estacionarias;

- Entonces, los coeficientes no cambian dentro de la muestra (validez interna);
- Y los resultados pueden extrapolarse fuera de la muestra (validez externa).
- Ésta es la otra cara (la de series temporales) de la correspondiente a la de i.i.d.

(b)  $(Y_t, X_t)$  y  $(Y_{t-j}, X_{t-j})$  se convierten en independientes cuando  $j$  tiende a infinito

- Periodos muy apartados en el tiempo pueden considerarse como si dieran lugar a experimentos separados.
- Con datos de sección cruzada, suponíamos que  $Y$ ,  $X$  eran i.i.d., como consecuencia del muestreo aleatorio simple – lo que conducía al TLC.
- Una versión del TLC con series temporales es la de independencia al aumentar la separación temporal – el supuesto 2(b) es la contrapartida de series temporales a la de i.i.d.

## Bajo los supuestos del modelo de retardos distribuidos:

- MCO proporciona estimadores consistentes de  $\beta_1$ ,  $\beta_2, \dots, \beta_r$  (los multiplicadores dinámicos)
- La distribución muestral de  $\hat{\beta}_1$ , etc., es normal
- *Sin embargo*, la fórmula para la varianza de la distribución muestral no es la usual (sección cruzada i.i.d.), porque  $u_t$  no es i.i.d. – está correlacionada serialmente.
- Esto significa que los errores estándar usuales por MCO están mal. En su lugar, utilizaremos *SEs* que sean robustos a la autocorrelación y heteroscedasticidad

# Errores estándar HAC

## (SW Sección 13.4)

- Cuando  $u_t$  está correlacionada serialmente, la varianza de la distribución muestral del MCO es diferente.
- Consecuentemente, necesitaremos utilizar una formula distinta para los errores estándar.
- Esto es fácil con STATA y otros (pero no todos) programas informáticos.

## *La parte matemática...*

Considere primero el caso sin retardos:

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

Recuerde que el MCO es:

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

entonces

$$\hat{\beta}_1 - \beta_1 = \frac{\frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})u_t}{\frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})^2} \quad (SW \text{ App. 4.3})$$

Es decir,

$$\hat{\beta}_1 - \beta_1 \cong \frac{\frac{1}{T} \sum_{t=1}^T v_t}{\sigma_X^2} \text{ en muestras grandes}$$

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

Por tanto,

$$\text{var}(\hat{\beta}_1) = \text{var}\left(\frac{1}{T} \sum_{t=1}^T v_t\right) / (\sigma_X^2)^2$$

*¿Qué sucede con datos de series temporales?*

Con  $T = 2$ :

$$\begin{aligned} \text{var}\left(\frac{1}{2} \sum_{t=1}^2 v_t\right) &= \text{var}\left[\frac{1}{2}(v_1 + v_2)\right] \\ &= \frac{1}{4}[\text{var}(v_1) + \text{var}(v_2) + 2\text{cov}(v_1, v_2)] \end{aligned}$$



$$= \frac{1}{2}\sigma_v^2 + \frac{1}{2}\rho_1\sigma_v^2 \quad (\rho_1 = \text{corr}(v_1, v_2))$$

$$= \frac{1}{2}\sigma_v^2 \times f_2, \text{ donde } f_2 = (1 + \rho_1)$$

- Con datos i.i.d.,  $\rho_1 = 0$  implica  $f_2 = 1$  – que nos da la fórmula usual de  $\text{var}(\hat{\beta}_1)$ .
- Con datos de series temporales,  $\rho_1 \neq 0$ ; por tanto, la fórmula de  $\text{var}(\hat{\beta}_1)$  **no** es la usual.
- Los SEs convencionales del MCO están mal cuando  $u_t$  está correlacionada serialmente.

## Expresión de la $\text{var}(\hat{\beta}_1)$ para un valor $T$ general

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

por tanto

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

donde

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

Los *SEs* están multiplicados por un factor  $f_T$  (¡que puede ser bastante grande!)

# HAC SEs

- *SEs* convencionales (consistentes a la heteroscedasticidad o no) están mal cuando existe autocorrelación.
- Necesitamos una nueva fórmula que produzca *SEs* robustos tanto a la heteroscedasticidad como a la autocorrelación.

## *Necesitamos HAC SEs*

- Si conociéramos  $f_T$ , podríamos hacer el ajuste.
- Pero no lo conocemos – depende de las autocorrelaciones desconocidas.
- HAC *SEs* sustituye  $f_T$  por un estimador de  $f_T$

## HAC SEs

$$\text{var}(\hat{\beta}_1) = \left[ \frac{1}{T} \frac{\sigma_v^2}{(\sigma_X^2)^2} \right] \times f_T, \text{ donde } f_T = 1 + 2 \sum_{j=1}^{T-1} \left( \frac{T-j}{T} \right) \rho_j$$

El estimador más utilizado es:

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

- $\hat{f}_t$  se denomina ponderación de Newey-West
- $\tilde{\rho}_j$  es un estimador de  $\rho_j$
- $m$  es un parámetro de *truncamiento*
- ¿Qué es un parámetro de truncamiento?
  - Utilice el método de Goldilocks
  - O, por ejemplo,  $m = 0.75T^{1/3}$

# Ejemplo: zumo de naranja

```
. 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 l1fdd if tin(1950m1,2000m12), r;   NOT HAC SEs
```

Regression with robust standard errors

```
Number of obs =      612
F( 1, 610) =      3.97
Prob > F      =      0.0467
R-squared     =      0.0101
Root MSE     =      5.0438
```

---

dlpoj	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
l1fdd	.1529217	.0767206	1.99	0.047	.0022532 .3035903
_cons	-.2097734	.2071122	-1.01	0.312	-.6165128 .196966

---

# Ahora calcule los *SEs* de Newey-West:

```
. newey dlpoj l1fdd if tin(1950m1,2000m12), lag(8);
```

```
Regression with Newey-West standard errors  
maximum lag : 8
```

```
Number of obs = 612  
F( 1, 610) = 3.83  
Prob > F = 0.0507
```

---

dlpoj	Coef.	Newey-West Std. Err.	t	P> t	[95% Conf. Interval]	
l1fdd	.1529217	.0781195	1.96	0.051	-.000494	.3063375
_cons	-.2097734	.2402217	-0.87	0.383	-.6815353	.2619885

---

*Uses autocorrelations up to  $m=8$  to compute the SEs*

*rule-of-thumb:  $0.75*(612^{1/3}) = 6.4 \approx 8$ , rounded up a little.*

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

## Ejemplo: zumo de naranja (cont.)

```
. global lfdd6 "fdd 11fdd 12fdd 13fdd 14fdd 15fdd 16fdd";  
. 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
```

---

dlpoj	Coef.	Newey-West Std. Err.	t	P> t	[95% Conf. Interval]
fdd	.4693121	.1359686	3.45	0.001	.2022834 .7363407
11fdd	.1430512	.0837047	1.71	0.088	-.0213364 .3074388
12fdd	.0564234	.0561724	1.00	0.316	-.0538936 .1667404
13fdd	.0722595	.0468776	1.54	0.124	-.0198033 .1643223
14fdd	.0343244	.0295141	1.16	0.245	-.0236383 .0922871
15fdd	.0468222	.0308791	1.52	0.130	-.0138212 .1074657
16fdd	.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)?

# ¿Necesitamos HAC *SEs* cuando estimamos un modelo AR or ADL? **NO.**

- El problema que resuelve los HAC *SEs* es el de la autocorrelación serial en  $u_t$
- Si  $u_t$  no está serialmente correlacionada, podremos seguir utilizando los *SEs* del MCO.
- En los modelos AR o ADL, no existe correlación serial si hemos incluido suficientes retardos de  $Y$ 
  - Si hemos incluido suficientes retardos de  $Y$ , no podremos estimar el término de error utilizando valores pasados de  $Y$ , o equivalentemente, valores pasados de  $u$  – consecuentemente  $u$  no está serialmente correlacionada.



# Estimación de efectos causales dinámicos con regresores estrictamente exógenos

## (SW Sección 13.5)

- $X$  es estrictamente exógena si
$$E(u_t | \dots, X_{t+1}, X_t, X_{t-1}, \dots) = 0$$
- Si  $X$  es estrictamente exógena, existen formas más eficientes para estimar efectos causales dinámicos que por medio del modelo de retardos distribuidos.
  - MCG
  - ADL
- Pero la condición de exogeneidad estricta es muy fuerte, poco plausible en la práctica.

# Análisis de los datos de zumo de naranja

## (SW Sección 13.6)

¿Cuál es el efecto causal dinámico (los *multiplicadores dinámicos*) de un incremento unitario en FDD sobre el precio del zumo?

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

- ¿Qué  $r$  usaremos?
  - ¿18? (Método de Goldilocks)
- ¿Qué  $m$  (*Parámetro de truncamiento de Newey-West*)?

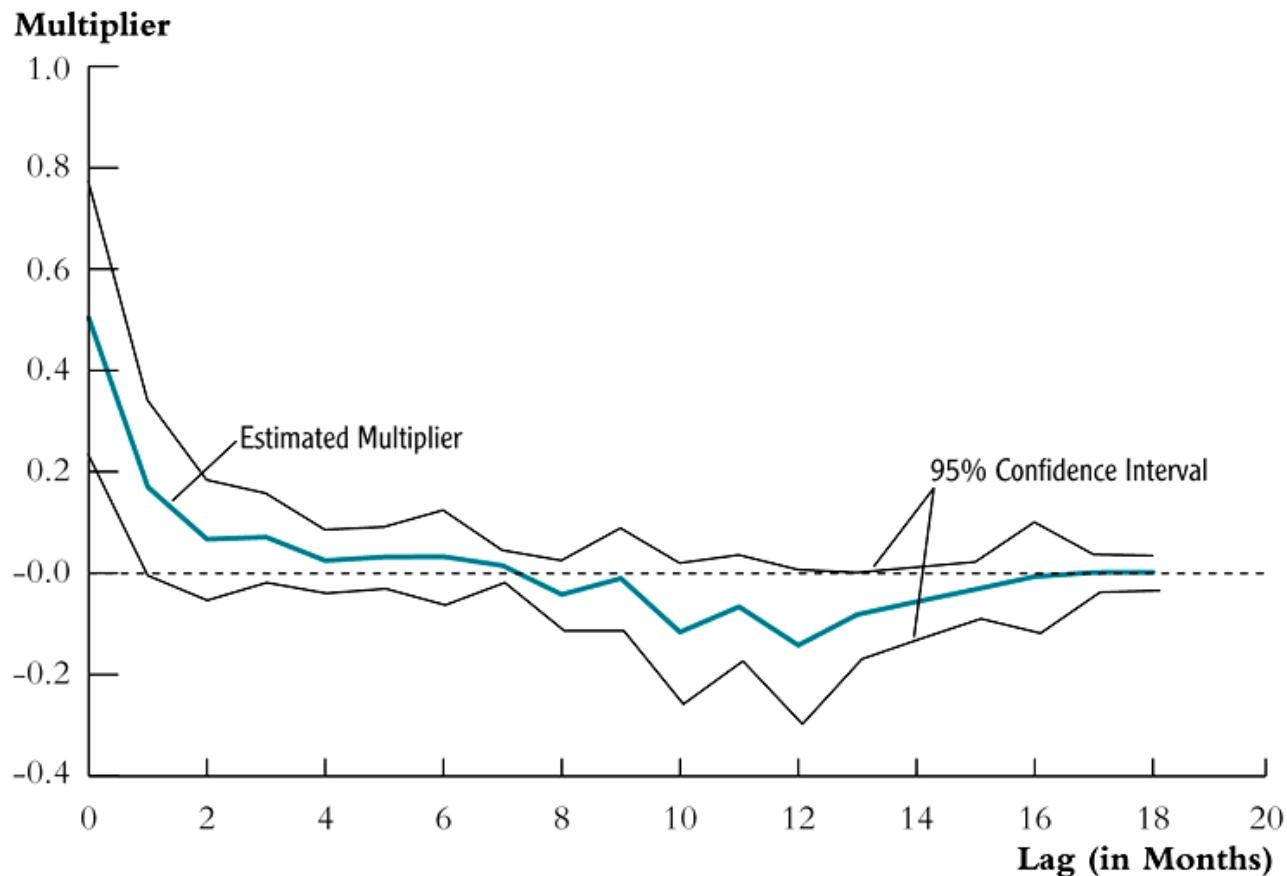
$$m = .75 \times 612^{1/3} = 6.4 \cong 7$$

**TABLE 13.1** 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

All regressions were estimated by OLS using monthly data (described in Appendix 13.1) from January 1950 to December 2000, for a total of  $T = 612$  monthly observations. The dependent variable is the monthly percentage change in the price of orange juice ( $\%ChgP_t$ ). Regression (1) is the distributed lag regression with the monthly number of freezing degree days and eighteen of its lagged values, that is,  $FDD_t, FDD_{t-1}, \dots, FDD_{t-18}$ , and the reported coefficients are the OLS estimates of the dynamic multipliers. The cumulative multipliers are the cumulative sum of estimated dynamic multipliers. All regressions include an intercept, which is not reported. Newey-West HAC standard errors, computed using the truncation number given in the final row, are reported in parentheses.

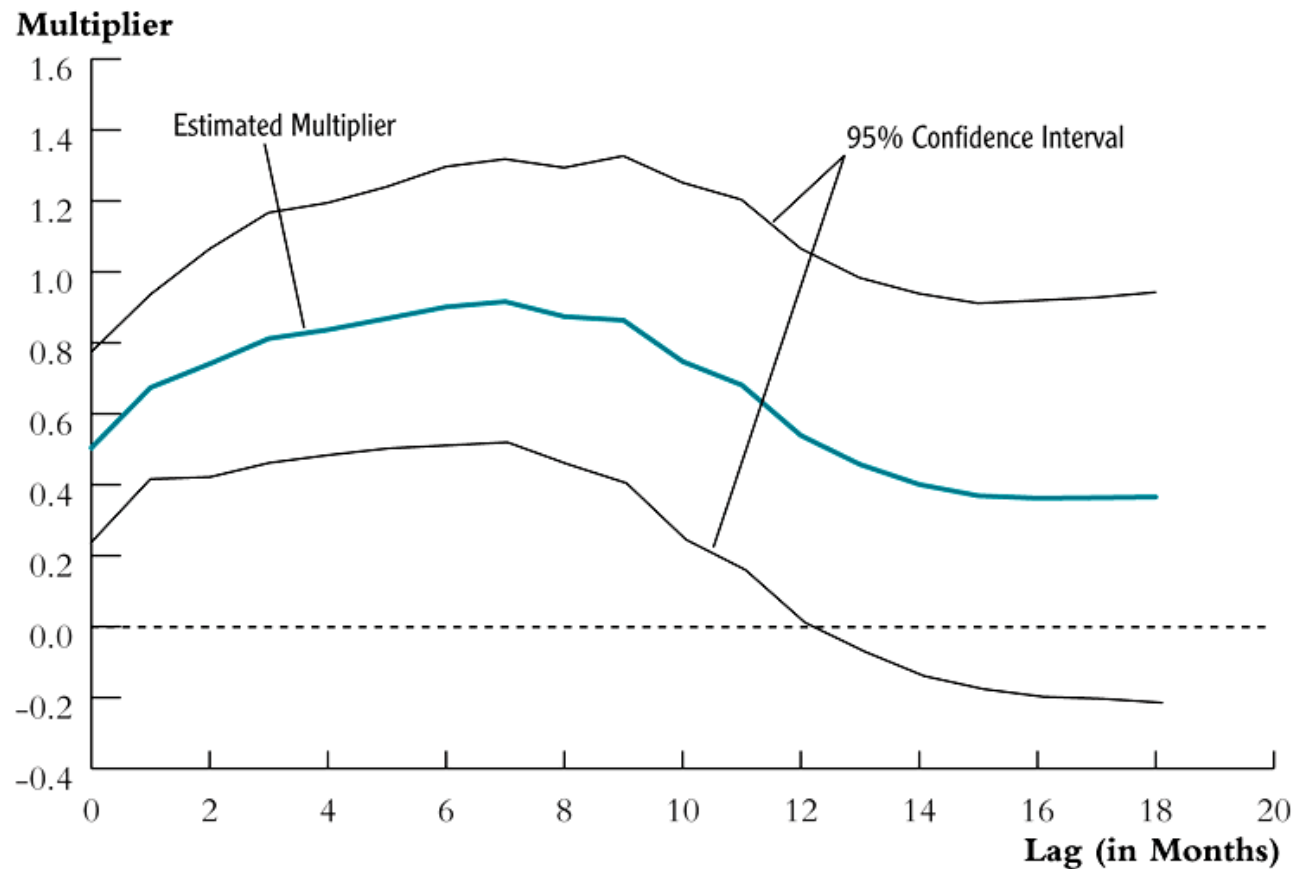
**FIGURE 13.2** The Dynamic Effect of a Freezing Degree Day (FDD) on the Price of Orange Juice



(a) Estimated Dynamic Multipliers and 95% Confidence Interval

The estimated dynamic multipliers show that a freeze leads to an immediate increase in prices. Future 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.

**FIGURE 13.2** The Dynamic Effect of a Freezing Degree Day (FDD) on the Price of Orange Juice

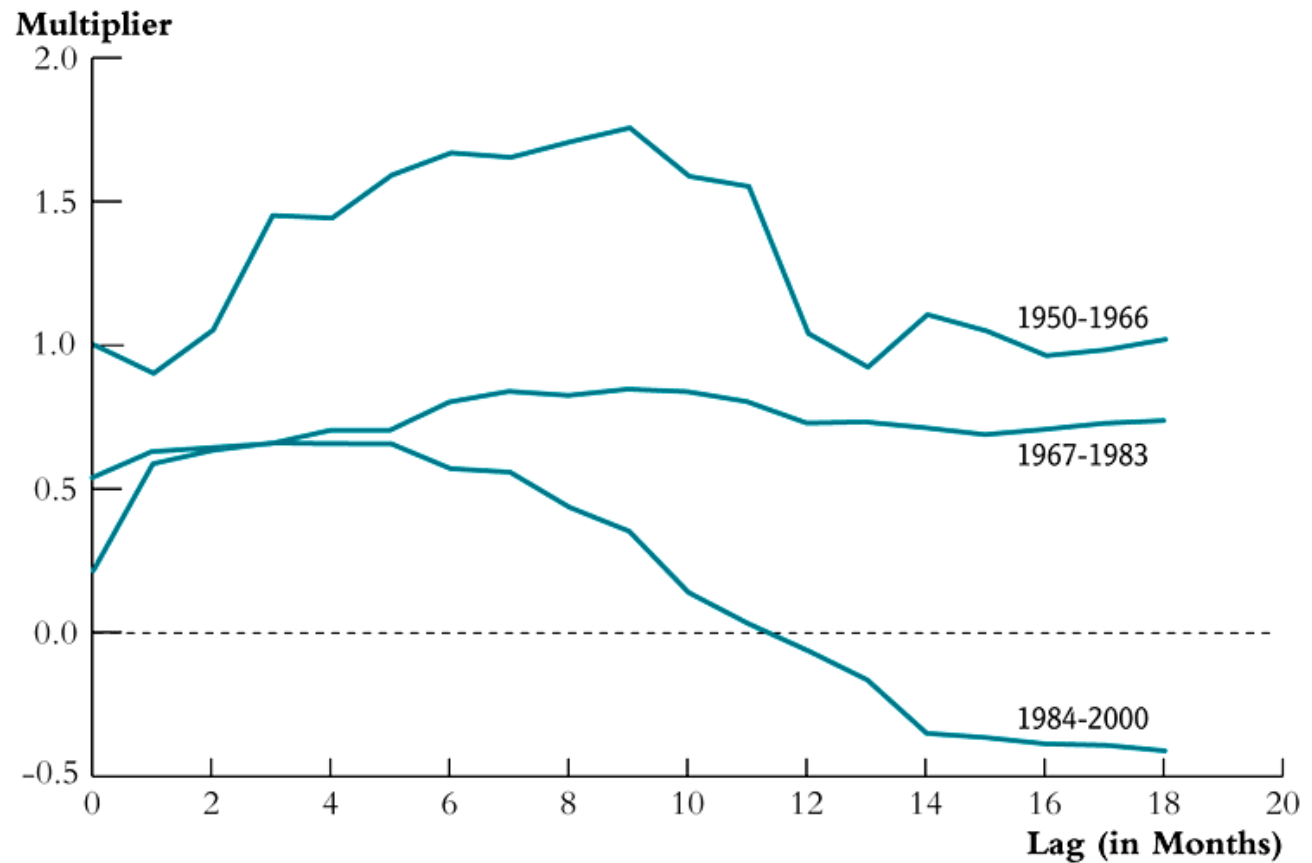


**(b)** Estimated Cumulative Dynamic Multipliers and 95% Confidence Interval

The estimated dynamic multipliers show that a freeze leads to an immediate increase in prices. Future 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.

**FIGURE 13.3** 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 twentieth century. A freeze had a larger impact on prices during 1950 to 1966 than later, and the effect of a freeze was less persistent during 1984–2000 than earlier.



Estos multiplicadores dinámicos han sido estimados por medio de un modelo de retardos distribuidos.

¿Obtendríamos estimadores más eficientes por medio de modelos de MCG o ADL?

- ¿Es  $FDD$  estrictamente exógena en la regresión de retardos distribuidos?

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

- Los comerciantes en el mercado del zumo no pueden cambiar el tiempo atmosférico.
- Ello implica que  $\text{corr}(u_t, FDD_{t+1}) = 0$

**¿Cuándo pueden ser estimados los efectos causales dinámicos? Es decir, ¿cuándo es plausible la hipótesis de exogeneidad? (SW Sección 13.7)**

En los siguientes ejemplos,

- ¿Es  $X$  exógena?
- ¿Es  $X$  estrictamente exógena?

**Ejemplos:**

1.  $Y$  = precios del zumo,  $X$  =  $FDD$  en Orlando
2.  $Y$  = exportaciones australianas,  $X$  = PIB de US (efecto de la renta norteamericana sobre la demanda de exportaciones australianas)



# Ejemplos

3.  $Y =$  exportaciones de la unión europea,  $X =$  PIB de US (efecto de la renta norteamericana sobre las exportaciones europeas)
4.  $Y =$  tasa de inflación de US,  $X =$  cambio porcentual en los precios mundiales del petróleo (OPEC) (efecto del precio de OPEC sobre el incremento de inflación)
5.  $Y =$  crecimiento del PIB,  $X =$  tipo de interés (efecto de la política monetaria sobre el crecimiento)
6.  $Y =$  cambio en la tasa de inflación,  $X =$  tasa de desempleo (la curva de Phillips)

# Exogeneidad

- Debe evaluar la exogeneidad y estricta exogeneidad caso por caso
- Exogeneidad no es a menudo plausible con datos de series temporales debido al sesgo de simultaneidad
- Exogeneidad estricta lo es aún menos debido a la retroalimentación de los datos de series temporales.

# Estimación de efectos causales dinámicos: Resumen

## (SW Sección 13.8)

- Los efectos causales dinámicos se pueden medir en teoría mediante experimentos controlados a través de mediciones repetidas a lo largo del tiempo.
- Cuando  $X$  es exógena, los estimaremos mediante una regresión de retardos distribuidos
- Si  $u$  está serialmente correlacionada, los  $SEs$  convencionales de MCO son incorrectos. Debemos utilizar los HAC  $SEs$
- ¡Es difícil decidir cuándo  $X$  es exógena o no!