تتضمن هذه العبارة أنه بمقارنة المعادلـة (3.60) بالمعادلة (3.61) ينتهي الباحثون إلـى التقدير نفسه للميل الحدي للاستهلاك، بينما سيحصل الباحث المستخدم للجدول (٣-٤أ) على حد ثابت أكبر مائة مرة عن ذلك الحد الثابت المشتق عن الجدول رقم (٣-٤ب). وكما رأينا فإن التقديرات لن تكون غير متناسقة لأن المتغيرات معرفـة بطريقة مختلفة.

وإثبات صحة هذه العـلاقـات مـن الـسـهـولـة بمـكـان. لاحـظ أولا أن $\overline{Y}_d = (1/100)y_d$ وأن $\overline{C} = (1/100)\overline{c}$. حينئذ

$$\hat{b}_t = \frac{\Sigma\left(Y_{dt} - \overline{Y}_d\right)C_t}{\Sigma\left(Y_{dt} - \overline{Y}_d\right)^2} = \frac{\Sigma\left(y_{dt} - \overline{y}_d\right)c_t}{\Sigma\left(y_{dt} - \overline{y}_d\right)^2} = \hat{b}_2 \tag{3.62}$$

حيث إن b_1 سيكون المقدر لـ b الذي يحصل عليه بوساطة الجدول رقم (٣-٤ب) والنموذج (3.61)، و b_2 يناظر الجدول رقم (٣-٤أ) والنموذج (3.60)وبالنسبة للحدود الثابتة يكون لدينا:

$$\hat{A} = \overline{C} - \hat{b}\overline{Y}_d = \frac{1}{100}\left(\overline{c} - \hat{b}\overline{y}_d\right) = \frac{1}{100}\hat{a} \tag{3.63}$$

دعنا الآن نعمّم نتائجنا، اعتبر النموذج:

$$y_t = a + bx_t + u_t \tag{3.64}$$

والآن دع:

$$Y_t = s_1 y_t, \quad X_t = s_2 x_t \tag{3.65}$$

حيث إن s_1 و s_2 ثوابت أو عوامل ترجيـح scale factors وبإحلال المعادلـة (3.65) محل المعادلة (3.64) تكون لدينا العلاقة التالية بين Y_t و X_t:

$$\begin{aligned} Y_t &= as_1 + \left(\frac{bs_1}{s_2}\right)X_t + s_1 u_t \\ &= A + BX_t + U_t \end{aligned} \tag{3.66}$$

حيث إن $A=s_1a$، $B=bs_1/s_2$ و $U_t = s_1u_t$. وهكذا، إذا قام أحد الباحثين بتقدير المعادلة (3.64)، بينما قام باحث آخر باختيار وحدات ترجيح أخرى (فأسقط، مثلا، الأصفار، غير الضرورية) لـY_t و X_t كما في المعادلة (3.65) ثم قام بتقدير المعادلة (3.66) بعد ذلك فإن العلاقة بين تقديرات معلماتها المناظرة يمكن الحصول عليها بوساطة:

$$\hat{A} = s_1\hat{a}, \quad \hat{B} = \hat{b}\left(\frac{s_1}{s_2}\right) \tag{3.67}$$

في ضوء المعادلة (3.67) تكون العلاقات بين تباينات المقدرات على النحو:

$$\sigma^2_{\hat{A}} = s_1^2\sigma^2_{\hat{a}}, \quad \sigma^2_{\hat{B}} = \left(\frac{s_1}{s_2}\right)^2 \sigma^2_{\hat{b}} \tag{3.68}$$

وعلى الرغم من أننا لن نقوم بإثباتها هنا، فإنه يمكن إثبات أن العلاقات بين مقدرات التباينات هي نفسها، بالضبط، كنظائرها في المعادلة (3.68):

$$\hat{\sigma}^2_{\hat{A}} = s_1^2\hat{\sigma}^2_{\hat{a}}, \quad \hat{\sigma}^2_{\hat{B}} = \left(\frac{s_1}{s_2}\right)^2 \hat{\sigma}^2_{\hat{b}} \tag{3.69}$$

فإذا اعطينا عوامل الترجيح، فإنه يمكن، عند ذلك، اشتقاق نتائج إحدى هذه الدراسات مباشرة من النتائج الأخرى.

قبل أن نعطي مثالا لنظرية الترجيح هذه، ربما ينبغي أن نشير إلى ما قد يكون واضحاً. إذا كانت لدينا افتراضات خاصة بالمعلمات a أو b في المعادلة (3.64) فإن هذه الافتراضات قد تختبر أما بدلالة المعادلة (3.64) أو المعادلة (3.66). فمثلا، نجد الافتراض $b = b^\circ$ يماثل الافتراض $B = b^\circ(s_1/s_2)$. ويمكن أن نذكر عبارات مشابهة للافتراضات المختصة بالمعلمة a. على الرغم من أننا لن نقوم بذلك هنا، فإنه يمكن اثبات أن قبول افتراض معين أو رفضه لـ a أو b والذي يختار بدلالة $\hat{a}$ أو $\hat{b}$ كما يشتق من المعادلة (3.64) فقط إذا كان الافتراض المناظر والمرتبط

بـ A أو B والذي يختار في المعادلة (3.66) قد قبل أو رفض.* وبمعنى آخر، فإن نسب t لـ $\hat{a}$ و$\hat{b}$ تكون متماثلة مع تلك لـ $\widehat{B}$ و $\widehat{A}$.

مثال

دعنا الآن نأخذ تطبيقا مبسطا لمبادئ الترجيح هذه. افترض أننا مهتمـون، مرة أخرى، بدالة الاستهلاك:

$$c_t = a + by_{dt} + u_t \tag{3.70}$$

افترض، أيضا أننا نجمع بيانات عــن c_t , y_{dt}، وأننا نقدر معلماتنا a و b، وأخيراً نختبر الافتراض $b = b_0$. افترض أننا قد علمنا أن البيانات التي استخدمناها غير دقيقة، وبالتحديد بسبب الطريقة التي جمعت بها البيانات، فإن أرقامنا عن الدخل المتاح تفوق البيانات بمقدار ١٠٪، ولكن، يفترض أن بياناتنا المرتـبـطـة بالاستهلاك دقيقة. ويصبح التساؤل حول ما إذا كان من المطلوب إعادة الدراسـة بالكامل باستخدام البيانات المصححة.

دع y^*_{dt} هو مقياسنا للدخل المتاح. يتضمن خطأ القياس أن:

$$y^*_{dt} = (1.1)y_{dt} \tag{3.71}$$

حيث إن y_{dt} هي القيمة الحقيقية للدخل المتاح. وبإحلال المعادلة (3.71) في دالتنا للاستهلاك بالمعادلة (3.70) نحصل على:

$$\begin{aligned} c_t &= a + \left(\frac{b}{1.1}\right)y^*_{dt} + u_t \\ &= a + By^*_{dt} + u_t \end{aligned} \tag{3.72}$$

* يمكن للقارئ أن يقنع نفسه بهذا عن طريق ملاحظة أنه، في ضوء المعادلة (3.69)

$$\frac{\hat{B}-B}{\hat{\sigma}_{\hat{B}}} \equiv \frac{\hat{b}-b}{\hat{\sigma}_{\hat{b}}} \quad \text{و} \quad \frac{\hat{A}-A}{\hat{\sigma}_{\hat{A}}} \equiv \frac{\hat{a}-a}{\hat{\sigma}_{\hat{a}}}$$

وهكذا، فإن فترات الثقة لـ A و B المبنية على المعادلة (3.66) هي، ببساطة، مرجحة لأعلى أو أسفل للفترات المناظرة كما تشتق من المعادلة (3.64).

حيث إن $B = (b/1.1)$.

المعادلة (3.72) هي النموذج المرتبط باستخدام بياناتنا غير الدقيقة. باستخدام نتائجنا بالمعادلة (3.67) يمكننا أن نرى من المعادلة (3.72) أن تقديرنا للحد الثابت a بالإضافة إلى نتائج اختبار الفرضية المرتبطة بذلك الحد لاتزال صحيحة. ولكن تقديرنا للميل الحدي للاستهلاك قد يكون منخفضا جدا لأن مقدرا غير متحيز لـ MPC يمكن أن يكون $\hat{b} = \hat{B}(1.1)$. وهكذا، ينبغي أن نضرب تقديرنا في (1.1) إضافة إلى ذلك، ينبغي علينا أن نعيد اختبار فرضيتنا المرتبطة بقيمة b. ومن السهل القيام بذلك إذا لاحظنا أن الفرضية (مثلاً $b = b_0$) تتضمن الفرضية $B=b_0/1.1$. وفي الحقيقة فإن قدرا ضئيلا من عملنا الأصلي ينبغي إعادة عمله.

وهناك ملاحظة أخيرة ترتبط بأهمية تقريب الأرقام العشرية (لعدد معقول) عند تقرير النتائج. وهذا لايسهل، فقط، المعادلة ولكن يجنبنا الدقة الوهمية أيضا، وربما تتذكر أنه، عندما قدرنا دالة الاستهلاك في الفصل الثاني، استخدمنا بيانات بوحدات من بلايين الدولارات دون كسور، وباستخدام هذه الأرقام، قدرنا معادلة الانحدار:

$$\hat{C} = 13 + 0.89Y_d$$

ونتيجة لاستخدام الحاسوب في الوصول إلى قيم هذه المعاملات، فيمكننا الحصول على معاملات مقدرة ذات عدد عشري كبير، لذلك يمكننا أن نكتب نتائج المعادلة السابقة في الشكل:

$$\hat{C} = 13.186537 + 0.889632Y_d$$

ولكن، من الواضح أن ذلك غير مطلوب بل غير مرغوب فيه أيضا. فطالما أن بياناتنا الأساسية صحيحة، فقط، لأقرب مليون من الدولارات، فإن من غير المعقول أن نحاول التنبؤ بمستوى الاستهلاك لأقرب ألف دولار. هذا يعطينا مايطلق عليه وهم الدقة، والذي يكون غير ممكن إذا استخدمنا البيانات الخام. ينبغي علينا، عند عرض النتائج، أن نضمن تناسب عدد الأرقام العشرية المهمة الموجودة في التقرير مع مستوى الدقة التي ستسمح بها البيانات الأساسية.

(٣-٤) استخدام المتغيرات المبطأة

اعتبرنا حتى الآن الشكل التالي من العلاقات:

$$C_t = a + bY_{dt} + u_t$$

$$\dot{W}_t = a + b\left(\frac{1}{R_t}\right) + u_t$$

حيث يشير الدليل السفلي t، في تحليل السلاسل الزمنية، إلى عنصر الزمن. وتشترك جميع هذه العلاقات في خاصية واحدة، في الأقل: قيمة المتغير التابع مرتبطة بقيمة المتغير المستقل عند النقطة نفسها من الزمن (أوعلى مدى الفترة الزمنية نفسها). فمثلا، افترضت نماذجنا أن الإنفاق الاستهلاكي في ١٩٥٠م يعتمد على الدخل المتاح في ١٩٥٠م.

ولكن، غالبا، مايتعامل الاقتصاديون مع نماذج لاتكون جميع المتغيرات فيها مرتبطة بالفترة الزمنية نفسها. افترض، مثلاً، أننا نحاول تفسير حجم الانفاق الاستهلاكي لمجموعة من الأفراد يحصلون على دخولهم في نهاية كل شهر. قد نتوقع أن يقوم هؤلاء الأفراد بانفاق نسبة من دخولهم خلال الشهر التالي. ولذا، تتوافر لدينا سلسلة زمنية يعتمد فيها الإنفاق في أحد الشهور على الدخل المتاح في الشهر السابق له. فإذا جعلنا t ترمز إلى الفترات بالشهور، يصبح لدينا:

$$C_t = a + bY_{d(t-1)} + u_t \qquad (3.73)$$

والتي قد تشير، على سبيل المثال، إلى:

$$C_{June} = a + bY_{d(May)} + u_{June}$$

وهكذا، فإن الإنفاق الاستهلاكي t يعتمد على الدخل الذي حصل عليه خلال الفترة (t-1) وتعبر عن ذلك بالقول إن الاستهلاك يتباطأ خلف الدخل بمقدار فترة زمنية واحدة، أو إن الاستهلاك يعتمد على Y_d مع فترة إبطاء واحدة.

وتوجد مجموعة أخرى من الافتراضات يمكن أن تؤدي إلى نموذج علـى غرار المعادلة (3.73). وتظهر هذه على النحو التالي. افترض أن:

$$C_t^p = a + bY_{dt}^e \qquad (3.74)$$

حيث:

C_t^p = حجم الإنفاق الاستهلاكي المخطط للفترة المقبلة t، وأن

Y_{dt}^e = الدخل المتوقع للفترة المقبلة t.

افترض، لتبسيط التحليل، أن:

$$Y_{dt}^c = Y_{d(t-1)} \qquad (3.75)$$

يبين هذا الافتراض أن الأفراد يتوقعون أن يتماثل دخلهم في الفترة المقبلـة t مع الدخل الذي حصلوا عليه في الفترة الحالية (t-1). إفترض، أيضا، أن:

$$C_t = C_t^p + u_t \qquad (3.76)$$

حيث ترمز C_t إلى الإنفاق الاستهلاكي الفعلي في الفترة t و u_t هو الخطأ العشوائي. أي أن الإنفاق الفعلي يختلف عن الإنفاق المخطط لوجود متغير عشوائي له قيمة متوسطة صفرية. لذا ففي المتوسط، يتعادل الإنفاق الفعلي مع الإنفاق المخطط. في هذه الحال تمثـل u تأثير الأحداث غير المتوقعة على حجم الإنفاق الفـعـلـي كفاتورة زيارة طبيب مثلا. وعلى أي حال، إذا قمنا بالتعويض عن C_t^p و Y_{dt}^e في المعادلة (3.74) نحصل على:

$$C_t = a + bY_{d(t-1)} + u_t \qquad (3.77)$$

والمتماثلة مع المعادلة (3.75).

وقد تفيد أنواع العلاقات المتباطئة في شرح السلوك الاستثماري والتغيرات في الأجور أيضا. فعلى سبيل المثال، لايتخذ قرار الاستثمار في الحال، وحتى إذا اتخذ هذا القرار في الحال فإن تنفيذه يستغرق وقتا. ولهذا السبـب، قـد ندعي أن:

$$I_{t-1}^{d} = a + br_{t-1} \tag{3.78}$$

فقرار الاستثمـار، I^d، في الفترة (t-1)، يعتمد على معدل الفـائـدة r، في تلك الفترة نفسها، ولكن افترض أن:

$$I_t = I_{t-1}^{d} + u_t \tag{3.79}$$

حيث يعتمد الإنفاق الاستثماري على قرارا الاستثمار في الفترة الزمنية السابقة (أو مع فترة إبطاء واحدة). وبالتعويض عن I_{t-1}^{d} في المعادلة (3.79). نحصل على علاقة تشبه العلاقة السابقة للاستهلاك أي:

$$I_t = a + br_{t-1} + u_t \tag{3.80}$$

ونترك للقاريء أن يثبت أنه، إذا دمجنا فترة إبطاء واحدة في منحنى فليبس حتى يصبح التغير بالنسبة المئوية في الأجور في الفتـرة t معتمدا على مستوى فائـض الطلب على العمل في الفترة السابقة(D_{t-1}^{*})، فسوف نحصل على:

$$\dot{W}_t = a + b\left(\frac{1}{R_{t-1}}\right) + u_t \tag{3.81}$$

هل وجود مثل هذا الإبطاء يبدو منطقيا ؟

والتساؤل الذي يثور الآن هو ما إذا كان من الممكن لنموذجنا أن يعـالـج مشكلة تقدير العلاقة التي تتضمن متغيرات مبطأة ؟ والإجابة هي نعم. وحتـى يمكننا رؤية ذلك اعتبر النموذج:

$$Y_t = a + bX_{t-} + u_t, \quad t = 1, 2, \cdots, n \tag{3.82}$$

حيث تتحقق الافتراضات الأساسية المرتبطة بالخطأ العشوائي u_t. ويتضح من هذه المعادلة أن Y تعتمد على X المتخلفة عنها بفترة إبطاء واحدة. افترض أنه يوجـد لدينا عدد n من المشاهدات عنY_t , X_t والتي يمكن أن نعبر عنها في الجدول رقم (٣-٥).

جدول رقم (٣-٥)

Y	X
Y_1	X_1
Y_2	X_2
.	.
.	.
Y_n	X_n

لاحظ أن Y_t ليست مرتبطة بـ X_t إنها تعتمد على X_{t-1}. ولهذا السبب، ينبغي أن نضع أزواج القيم من Y وX بحيث نجعل مقابل كل قيمة من قيم Y قيمة X في الفترة السابقة لها كما في الجدول رقم (٣-٦).

جدول رقم (٣-٦)

Y	X
Y_1	X_0
Y_2	X_1
Y_3	X_2
.	.
.	.
Y_n	X_{n-1}

إذا أردنا أن نضع القيم المشاهدة لـ Y وX في شكل انتشار، فإن كل نقطة في الشكل سوف تمثل قيمة Y وقيمة X في الفترة السابقة. وهذه النقاط هي التي نوفق بها خطنا للانحدار.

لاحظ أننا، عند الانتقال من الجدول رقم (٣-٥) إلى الجدول رقم (٣-٦) خسرنا مشاهدة واحدة، حيث لن يكون بامكاننا استخدام Y_1 طالما لاتوجد لدينا مشاهدة X_0. وبالمثل، فلن يمكننا، أيضا، استخدام X_n طالما أننا لانعرف Y_{n+1}. وهكذا يقلل النموذج هذا فترة الإبطاء الواحدة من حجم عينتنا بمقدار مشاهدة واحدة إلى (n-1). أي يوجد في الجدول رقم (٣-٦) عدد (n-1)، فقط، من المشاهدات الزوجية التي يمكن استخدامها في نموذجنا.

ولتقدير علاقتنا المبطأة هذه، نوجد متغيرًا جديدًا هو $Z_t = X_{t-1}$ (وهكذا فإن قيمة Z في أي فترة زمنية تتساوى، ببساطة، مع قيمة X في الفترة السابقة لها). ولذا يمكن إعادة كتابة نموذجنا الأساسي في المعادلة (3.82) على النحو:

$$Y_t = a + bZ_t + u_t, \quad t = 2,\cdots,n. \tag{3.83}$$

وتصبح مقدراتنا لـ a , b هي:

$$\hat{b} = \frac{\sum_{t=2}^{n}\left(Z_t - \overline{Z}\right)Y_t}{\sum_{t=2}^{n}\left(Z_t - \overline{Z}\right)^2} \tag{3.84}$$

$$\hat{a} = \overline{Y} - \hat{b}\overline{Z}$$

حيث:

$$\overline{Y} = \frac{\sum_{t=2}^{n} Y_t}{n-1}, \quad \overline{Z} = \frac{\sum_{t=2}^{n} Z_t}{n-1},$$

لاحظ أننا، في هذه الحسابات، نسقط كلاً من Y_1 و X_n.

مثال

لتوضيح هذه الطريقة، دعنا نعود إلى دالة الاستهلاك التي قدرناها في الفصل الثاني والتي تأخذ الشكل:

$$C_t = a + bY_{dt} + u_t$$

دعنا الآن نقدر هذه الدالة مع فترة إبطاء واحدة.

$$C_t = a + bY_{d(t-1)} + u_t \qquad (3.85)$$

في المعادلة (3.85)، نفترض أن الاستهلاك في أي سنة معينة يعتمد على مستوى الدخل المتاح في السنة السابقة.

بالعودة إلى الجدول رقم (٢-٢)، ومزاوجة القيم المشاهدة للاستهلاك مع الدخل المتاح في السنة السابقة لها، نحصل على الجدول رقم (٣-٧).

جدول رقم (٣-٧)

السنة	الاستهلاك ببلايين الدولارات	السنة	الدخلا المتاح ببلايين الدولارات
١٩٦١	٣٣٥	١٩٦٠	٣٥٠
١٩٦٢	٣٥٥	١٩٦١	٣٦٤
١٩٦٣	٣٧٥	١٩٦٢	٣٨٥
١٩٦٤	٤٠١	١٩٦٣	٤٠٥
١٩٦٥	٤٣٣	١٩٦٤	٤٣٨
١٩٦٦	٤٦٦	١٩٦٥	٤٧٣
١٩٦٧	٤٩٢	١٩٦٦	٥١٢
١٩٦٨	٥٣٧	١٩٦٧	٥٤٧
١٩٦٩	٥٧٦	١٩٦٨	٥٩٠

لاحظ أنه يتوافر لدينا الآن تسع مشاهدات، فقط، بدلا من عشر. وبتطبيق منهجنا في التقدير، نحصل على دالة الاستهلاك المتباطئة:

$$\hat{C} = \underset{(0.2)}{-20} + \underset{(44.3)}{0.98}Y_{d(t-1)}, \quad \underset{R^2=0.99}{n=9} \qquad (3.86)$$

حيث تعبر الأعداد الموجودة بين الأقواس عن القيم المطلقة لنسب t المناظرة. من النتائج نرى أن للمعادلة (3.86) قوة تفسيرية عالية، تماما كما هو الحال بالنسبة لدالة الاستهلاك التي قدرناها في الفصل الثاني، إلا أن الحد الثابت في معادلة الاستهلاك

المبطئة لايختلف معنويا عن الصفر (نسبة t تعادل 0.2 فقط) عند مستوى ثقة 95%. إضافة إلى ذلك، فإن الميل الحدي للاستهلاك MPC أعلى بدرجة كبيرة (0.98 مقابل 0.89). نتيجة لذلك، فإن السياسة الاقتصادية المبنية على دالة الاستهلاك المبطئة قد تختلف عن تلك المؤسسة على دالة غير مبطئة. من الواضح أننا نحتاج إلى طريقة تمكننا من التمييز بين هذين النموذجين. سنوجد مثل هذه الطريقة في الفصل الخامس، إضافة إلى ذلك، سوف نعالج انذاك نماذج أكثر عمومية للعلاقات المبطئة.

(٣-٥) التنبؤ Prediction

نتجه في هذا الجزء إلى التطبيق الثاني الأساسي لتحليل الانحدار. في الأجزاء الأولى من هذا الفصل، أوضحنا أن نتائج الانحدار يمكن استخدامها في اختبار الفرضيات حول السلوك الاقتصادي لمختلف الوحدات الاقتصادية. كما نستخدم بدرجة الأهمية نفسها، معادلات الانحدار المقدرة في التنبؤ بتأثير أحداث معينة على المتغيرات الاقتصادية. ربما تتذكر أننا ناقشنا في مقدمة الفصل الأول مشكلة المستشار الاقتصادي الذي يعمل على تقويم تأثير التخفيضات الضريبية (من أحجام مختلفة) على مستوى الإنفاق الاستثماري. فإذا افترضنا أن هذا المستشار يعلم، مثلا، المقدار الذي يتزايد به الدخل المتاح لكل تخفيض ضريبي، فإنه يمكنه في هذه الحال استخدام العلاقة المقدرة بين C و Y_d في التنبؤ بتأثير مختلف التخفيضات الضريبية على حجم الانفاق الاستهلاكي. وهنا يصبح لمعادلات الانحدار التي قدرناها دور مساعد حقيقي في تقويم الآثار الممكنة للسياسات الاقتصادية. ومن ثم، فإن تحليل الانحدار يمكن أن يساعد في فهم كمي لكيفية عمل الاقتصاد القومي، ومن ثم، التنبؤ بتأثير مختلف الاختيارات المتاحة لمصممي السياسة الاقتصادية.

والآن نرغب في فحص مشكلة التنبؤ فحصا أكثر دقة، افترض أن لدينا العلاقة ذات الشكل المعروف التالي:

$$Y_t = a + bX_t + u_t \qquad (3.87)$$

افترض، مثلا، أننا نعلم أصلا أن قيمــة X ستكون في فترة زمنية مستقبليــة X_f. فمثلاً إذا كانت X هي مستوى الدخل المتاح، قد نفترض أن X_f هي قيمة X التي ستنتج من تخفيض ضريبي معين. ومشكلتنا تتمثل في التنبؤ بقيم Y (أو Y_f) التي تناظر تلك القيمة المحــددة X_f. لاحظ أن Y_f هي القيمة المستقبلية لـ Y والمناظــرة لقيمة X_f المعطاة لـ X.

أول شئ ينبغي ملاحظته بالنسبة للتنبؤ هو أن Y_f ذاتها متغير عشوائي. فعلى سبيل المثال، وطبقا لنموذجنا في المعادلة (3.87) لدينا:

$$Y_f = a + bX_f + u_f \qquad (3.88)$$

حيث إن u_f هي قيمة الخطأ العشوائي في هذه الفترة المستقبليــة. والآن، وفـقــا لتحديداتنا المعيارية، لايمكننا التنبؤ بـ u_f بحكم أنه متغير عشوائي ليس مرتبطا بأي من القيم السابقة للخطأ العشوائي، أو بقيم المتغير المستقل X. وحتى إذا كنا نعلم a و b وبالتالي يمكننا حساب (a+bX_f)، فإننا لانزال غير قادرين على التنبؤ بــ Y بصورة دقيقة وذلك بسبب التأثير الذي لايمكن التنبؤ به لـ u_f.

إضافة إلى ذلك، سيكون هناك مصدر آخر لعدم التأكد أو عدم الدقة فــي تنبؤنا حيث إننا، عموما لانعرف أيا من a و b ولذا علينا أن نستخدم تقديرات لها حساب الجزء الأول من Y_f في المعادلة (3.88) أي قيمتها المتوسطة المناظرة لـ X_f:

$$Y_f^m = a + bX_f \qquad (3.88)$$

باختصار، سيكون هناك مصدران مختلفان للخطأ في تنبؤنا: الأثر الذي لايمكن التنبؤ به للخطأ العشوائي، u_F، واستخدام القيمة المقدرة للمعلمات a و b .

ولقد رأينا في المعادلــة (3.88)، وبافتراض أن قيمــة X_f معطاة، أن القيمــة المستقبلية المناظرة لــ Y، Y_f، هي متغير عشوائي. ولذا سيكون من المرغوب فيــه الحصول ليس، فقط، على تقدير النقطة والتنبؤ بــ Y_f، ولكن بناء فترة الثقة لها أيضا. أي أننا نرغب في الحصول على مقياس ما لدقة تنبؤنا.

سنوضح الآن طريقة لاستخدام نتائج الانحدار لتكوين التنبؤات وبناء فترات الثقة. سوف نقدم هذه الطريقة في خطوتين. وطالما أننا لانعرف أيا من a أو b، فإننا لانعرف Y_t^m. سنتجه أولا لاشتقاق مقدر لـY_f^m وآخر لتباين هذا المقدر. بعد الوصول إلى هذه المقدرات ستكون الخطوة الثانية التنبؤ بـY_f ذاتها. وتحديد فترات الثقة المرتبطة بها.

تقدير Y_f^m

لدينا من المعادلة (3.89) الصيغة التالية للقيمة المتوسطة لـY_f والمناظرة لـX_f وهي:

$$Y_f^m = a + bX_f \tag{3.89}$$

افترض الآن أن لدينا المقدرات $\hat{a}$ و $\hat{b}$ المؤسسة على عينة من Y و X ذات حجم n للفترات $t = 1,2,...,n$، وطالما أن f هي فترة مستقبلية فإن $n < f$ يصبح كل من $\hat{a}$ و $\hat{b}$ في ظل افتراضاتنا المعتادة، مقدرات غير متحيزة. ويترتب على ذلك أنه يمكننا استخدام الصيغة $\left(\hat{Y}_f = \hat{a} + \hat{b}X_f\right)$ مقدرا غير متحيز لـY_f^m، طالما أن لدينا، وبافتراض قيمة X معطاة كمايلي:

$$\begin{aligned} E\left(\hat{Y}_f\right) &= E\left(\hat{a} + \hat{b}X_f\right) = E(\hat{a}) + \left[E\left(\hat{b}\right)\right]X_f \\ &= a + bX_f = Y_f^m \end{aligned} \tag{3.90}$$

وبالعودة، مثلاً، إلى دالتنا المقدرة للاستهلاك في الفصل الثاني، افترض أن التخفيض الضريبي المقترح يرتبط بمستوى دخل متاح قدره 500 بليوناً من الدولارات. علينا حينئذ تقدير القيمة المتوقعة للإنفاق الاستهلاكي المرتبط بهذا التخفيض الضريبي على النحو:

$$\hat{C}_f^m = 13 + 0.89(500) = 13 + 445 = 458 \tag{3.91}$$

نلاحظ أن $\hat{Y}_f$ هو تقدير النقطة للقيمة المتوسطة لـ Y_f، Y_f^m، والمناظرة لـ

X_f. فإذا أردنا الحصول على فترة ثقة، أو اختبار فرضيات، Y_f^m فإننا نحتاج إلى توزيع احتمالي (أو دالة) لـ $\hat{Y}_f$. يمكن اشتقاق هذا التوزيع، بسهولة، من نظرية أساسية في الإحصاء، تصرح بأن التوليفات الخطية من المتغيرات الطبيعية هي ذاتها موزعة توزيعا طبيعيا. وطالما افترضنا طبيعية توزيع الأخطاء العشوائية فـإن كـلا من $\hat{a}$, $\hat{b}$ يكون موزعا توزيعا طبيعيا أيضا. وبافتراض أن قيمـة X_f معطاة، فـإن Y_f تصبح مؤلفا خطيا من $\hat{a}$ و $\hat{b}$ ومن ثم ينبغي أن تكون متغيرا عشوائيا موزعا توزيعا طبيعيا أيضا. أما القيمة المتوسطة لـ $\hat{Y}_f$ فهي Y_f. إضافة إلى ذلك، يمكن إثبات أن تباين $\hat{Y}_f$ هو*:

$$\sigma_{\hat{Y}_f}^2 = \sigma_u^2\left[\frac{1}{n} + \frac{\left(X_f - \overline{X}\right)^2}{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2}\right] \tag{3.92}$$

حيث إن $(\overline{X} = \sum_{t=1}^{n} X_t/n)$، وإن $X_1, ..., X_n$ هي المشاهدات لـ X التي بنيت عليها المقدرات $\hat{a}$ و $\hat{b}$ وبالتالي تكون $\hat{Y}_f$ موزعة $N\left(a + bX_f, \sigma_{\hat{Y}_f}^2\right)$.

وقبل أن نستمر في التحليل، علينا أن نلاحظ أن تباين $\hat{Y}_f$ يزداد مع مربع الفرق $(X_f - \overline{X})$، ويعني هذا إنه، كلما ابتعدت القيمة المعينـة لـــ X_f عن القيمـة المتوسطة للعينة من المشاهدات عـن X (والتي استخدمت لبناء مقدراتنـا $\hat{a}$ و $\hat{b}$) يتزايد تباين مقدرنا $\hat{Y}_f$، ويبدو هذا بدهيا حيث إنه، كلما ابتعدت القيمة المتنبأ بها للمتغير المستقل عن تلك القيم التي تقع ضمن مدى خبرتنا المشاهدة، انخفضت ثقتنا في دقة هذه التنبؤات. وقد نشعر بدرجة عالية من الثقة في تقدير القيـمـة

* لمعرفة كيفية الوصول إلى الصيغ الموجودة في هذا الجزء، انظر:

J. Johnston. *Econometric Methods*, 2nd ed. (New York: McGraw-Hill, 1972), pp. 38-43.

المتوسطة للإنفاق الاستهلاكي المرتبطة بمستوى الدخل المتاح القريب جدا من المستويات السائدة في السنوات الأخيرة. وعلى العكس من ذلك قد نشعر بدرجة كبيرة من عدم التأكد حول تقدير المستوى المتوقع من الإنفاق الاستهلاكي المرتبط بمستوى من الدخل المتاح يبلغ، بالتقريب، ضعف مستويات الدخل الحديثة.

وأخيرا، ينبغي أن نلاحظ أنه، طالما أن σ_u^2 غير معلومة عموما فإن تباين $\hat{Y}_f$ سيكون غير معلوم، ولذا، ينبغي تقديره. ومن الواضح أن المقدر المقترح هو:

$$\hat{\sigma}_{\hat{Y}_f}^2 = \hat{\sigma}_u^2 \left[\frac{1}{n} + \frac{\left(X_f - \overline{X}\right)^2}{\sum_{t=1}^{n} \left(X_t - \overline{X}\right)^2} \right] \tag{3.93}$$

الذي يتبين من مناقشتنا السابقة أنه غير متحيز.

$$E\left(\hat{\sigma}_{\hat{Y}_f}^2\right) = \sigma_{\hat{Y}_f}^2 \tag{3.94}$$

وتوضح المناقشة أعلاه أنه، إذا كان σ_u^2 معلوما، فإننا نستطيع الحصول على فترات الثقة، ونختبر الفروض المرتبطة بـ Y_f^m بملاحظة أن:

$$\frac{\left(\hat{Y}_f - Y_f^m\right)}{\sigma_{\hat{Y}_f}} \tag{3.95}$$

هو $N(0,1)$.

أما إذا كانت σ_u^2 غير معلومة، فإننا نلاحظ أن:

$$\frac{\hat{Y}_f - Y_f^m}{\hat{\sigma}_{\hat{Y}_f}} \tag{3.96}$$

هو متغير t بدرجات حرية تعادل (n-2).

على سبيل المثال، إذا كانت σ_u^2 غير معلومة، تصبح فترة ثقة 95% لـ Y_f^m:

$$\left(\hat{Y}_f \pm t_{n-2;0.975}\hat{\sigma}_{\hat{Y}f}\right) \tag{3.97}$$

وبالعودة إلى توضيحنا السابق، لدينا $\hat{C}_f = 458$ تناظر $Y_d = 500$. ولتحديد فترة ثقة 95% لـ C_f^m، فإنه:

$$458 \pm 2.31\left[3.4\sqrt{\frac{1}{10}+\frac{(500-469)^2}{85.810}}\right] = 458 \pm 2$$

التنبؤ بـ Y_f

نتجه الآن إلى القضية ذات الأهمية الرئيسية: التنبؤ بـ Y_f ذاتها وتحديد فترات الثقة المرتبطة بها. نلاحظ أولا، على افتراض أن قيمة X_f معطاة، أن مقدرنا (أو تنبؤنا) بالمستوى المستقبل لـ Y، Y_f، يكون متطابقا مع مقدرنا Y_f^m وتحديدا، $\hat{Y}_f = \hat{a} + \hat{b}X_f$ ويتبع هذا أن المكون العشوائي الذي لايمكن التنبؤ به في Y_f [انظر المعادلة (3.82)] له قيمة متوسطة صفرية. وبمعنى آخر فسوف نتنبأ بمستوى Y_f ببساطة عن طريق التنبؤ بقيمته المتوسطة.

في هذه الحال، يكون الخطأ في تنبؤنا هو:

$$e_f = Y_f - \hat{Y}_f \tag{3.98}$$

ويطلق على e_f خطأ التنبؤ The forecast error. وينتج عن افتراضاتنا أن خطأ التنبؤ له قيمة متوسطة صفرية:

$$\begin{aligned} E(e_f) &= E(Y_f) - E(\hat{Y}_f) \\ &= a + bX_f - a - bX_f = 0 \end{aligned} \tag{3.99}$$

افترض أن u_f، مثل u_1، ...، u_n، موزع توزيعا طبيعيا بقيمة متوسطة صفرية. وتباين σ_u^2. يترتب على ذلك، ومع افتراض أن قيمة X_f معطاة، أن Y_f موزع توزيعا طبيعيا، أيضا وأن قيمته المتوسطة هي $(a + bX_f)$، وأن تباينه هو σ_u^2. وطالما أن خطأ التنبؤ e_f في المعادلة (3.98) مؤلف خطيا من المتغيرات الطبيعية (تذكر أن $\hat{Y}_f$

متغير طبيعي)، إذن، ينبغي أن يكون هذا الخطأ، أيضا متغيرا طبيعيا.

والآن، وقد حددنا أن القيمة المتوسطة e_f هي الصفر، يمكننا أن نحدد تباين e_f عن طريق ملاحظة أن Y_f و $\hat{Y}_f$، مع افتراض أن قيمة X_f معطاة، مستقلان . على سبيل المثال (من المعادلة 3.88) نجد أن الخطأ العشوائي الوحيد الذي يعتمد عليه Y_f هو u_f. ولكن $\hat{Y}_f$ تعتمد على الأخطاء العشوائية u_1، ...، u_n لأن $\hat{a}$ و $\hat{b}$ قد بنيت، فقط، بدلالة المشاهدات المشتركة (Y_1, X_1)، (Y_2, X_2)، ... و (Y_n, X_n). وبفرض أن كلا خطأ عشوائي مستقل عن جميع الأخطاء العشوائية الأخرى فسيكون كل من Y_f و $\hat{Y}_f$ مستقلا، بافتراض أن قيمة X_f معطاة. ومن المعادلة (3.98) نجد أن e_f هي مولف خطي من متغيرين عشوائيين، تباينه هو (انظر ملحق الفصل الثاني):

$$\begin{aligned}\sigma_e^2 &= \sigma_{Y_f}^2 + \sigma_{\hat{Y}_f}^2 \\ &= \sigma_u^2 + \sigma_u^2\left[\frac{1}{n} + \frac{\left(X_f - \overline{X}\right)^2}{\Sigma\left(X_i - \overline{X}\right)^2}\right] \\ &= \sigma_u^2\left[1 + \frac{1}{n} + \frac{\left(X_f - \overline{X}\right)^2}{\Sigma\left(X_i - \overline{X}\right)^2}\right]\end{aligned} \tag{3.100}$$

وباختصار، يكون $e_f \sim N(0, \sigma_e^2)$.

وعلى نحو مشابه للتحليل السابق، نجد أن مقدراً غير متحيز لـ σ_e^2 يمكن أن يتخذ الشكل:

$$\hat{\sigma}_e^2 = \hat{\sigma}_u^2\left[1 + \frac{1}{n} + \frac{\left(X_f - \overline{X}\right)^2}{\Sigma\left(X_i - \overline{X}\right)^2}\right] \tag{3.101}$$

ونبني فترات الثقة (والتي يطلق عليها، أحيانا، فترات التنبـؤ) لـــ Y_f عن طريق ملاحظة أن:

$$\frac{e_f}{\sigma_e} = \frac{Y_f - \hat{Y}_f}{\sigma_e} \tag{3.102}$$

موزعٌ توزيعا طبيعيا معياريا $N(0,1)$ أو، في حالة عدم معرفة σ_u^2 أن:

$$\frac{e_f}{\hat{\sigma}_e} = \frac{Y_f - \hat{Y}_f}{\hat{\sigma}_e} \tag{3.103}$$

هو متغير t بدرجات حرية قدرها $n-2$. فعلى سبيل المثال، إذا كانت σ_u^2 غير معلومة فإن فترة الثقة 95% لـ Y_f ستكون:

$$\left(\hat{Y}_f \pm t_{n-2;0.975}\hat{\sigma}_e\right) \tag{3.104}$$

من المعادلــة (3.100) لاحظ أن $\sigma_e^2 > \sigma_{\hat{Y}_f}^2$ ومن المعادلتــين (3.101) و (3.93) أن $\hat{\sigma}_e^2 > \hat{\sigma}_{Y_f}^2$، ويترتب على هذه النتائج أن فترة الثقــة لـــ Y_f ستكون أوسع من الفترة لمستوى الثقة نفسه لـ Y_f^m [قارن المعادلة (3.104) بـ المعادلة (3.97)]. وهذا هو ماينبغي أن يكون. هناك صعوبتان في التنبؤ بـY_f، الأولى هي أن المعلمتين a و b غير معلومتين، والثانية هي أن الخطأ العشوائي u_f لايمكن التنبؤ به. أما عند التنبؤ بـ Y_f^m فتواجهنا صعوبة واحدة هي أن a و b غير معلومتين.

وعلى سبيل توضيح أخير، نعود، مرة أخرى، إلى معادلـتـنـا المـقـدرة للاستهلاك، ونحسب 95% فترة ثقة لمستوى الاستهلاك. افترض أن مستوى الدخل (كما سبق) هو $Y=500$ حينئذ، تكون فترتنا:

$$\left(\hat{a} + \hat{b}Y_d\right) \pm t_{n-2;0.975}$$

$$= 458 \pm 2.31\left[3.4\sqrt{1+\frac{1}{10}+\frac{(500-469)^2}{85,810}}\right] = 458 \pm 8$$

وكما لاحظنا، نجد أن فترة الثقة لـ C_f التي هي (458 ± 8) أوسع من تلك الخاصة بـ C_f^m التي هي (458 ± 2).

(٣-٦) مثال: التقدير لمنحنى طلب

نختم معالجتنا لنموذج الانحدار البسيط بتمرين توضيحي يتضمن تقديراً لمنحنى طلب. يعرض الجدول رقم (٣-٨) بعض البيانات الفعلية حول المبيعات السنوية وأسعار الدواجن في الولايات المتحدة الامريكية. يشير الجدول، بخاصة، إلى الاستهلاك السنوي للفرد وأسعار الاستهلاك السنوي مكمشة بالرقم القياسي لأسعار المستهلكين خلال السنوات من ١٩٤٨م إلى ١٩٦٣م.

سوف نستخدم هذه البيانات لتقدير منحنى طلب على الدواجن. افترض العلاقة التالية:

$$Q_t = aP_t^b e^{u_t}, \qquad (3.105)$$

حيث Q_t هي الاستهلاك الفردي بالرطل في الزمن t (خلال السنة t)، P_t هو السعر المناظر للدواجن مقاساً بالسنت Cent لكل رطل و u_t الخطأ العشوائي. على سبيل توضيح المقاييس المتضمنة، فإن قيمة $Q_t = 28.9$، تعني أنه، خلال السنة t كان الاستهلاك الفردي من الدجاج 28.9 رطل، وقيمة أخرى $P_t = 41.4$ تعني أن السعر المتوسط المناظر خلال تلك السنة كان 41.4 سنتا لكل رطل. يشير الجدول رقم (٣-٨) إلى أن هذه الأرقام تناظر عام ١٩٥٩م.

افترض الآن أن الخطأ العشوائي u_t يحقق الافتراضات كافة لنموذجنا المعتاد، حينئذ وكما أشرنا من قبل فيما يتصل بالمعادلة (3.43)، يمكن تفسير المعلمة b في النموذج [كما في المعادلة (3.105)] بأنها مرونة. وفي هذه الحال، تصف b النسبة المئوية المتوقعة للتغير في الاستهلاك الفردي السنوي من الدجاج المصاحب لمعدل تغير في السعر قدره ١٪.

وكما رأينا من قبل في هذا الفصل، يمكننا استخدام التحويلة اللوغاريتمية لجعل النموذج (3.105) يأخذ الشكل الخطي:

$$\ln\left(Q_t\right) = A + b\ \ln\left(P_t\right) + u_t \qquad (3.106)$$

حيث $A = \ln(a)$، وأخذت اللوغاريتمات للأساس e. وهذا يقترح، كما نوقش من قبل، أننا نأخذ، ببساطة، اللوغاريتم الطبيعي لكل من المتغيرين ثم نجري انحدارا للوغاريتم المتغير التابع $[\ln(Q_t)]$ على لوغاريتم المتغير المستقل $[\ln(P_t)]$، لكي نحصل على تقديرات لـ A و b. (مثلا، $\hat{A}$ و $\hat{b}$). حينئذ، يكون تقديرنا (المتحيز) لـ a هو $\hat{A}$.

جدول رقم (٣-٨) الاستهلاك الفردي والسعر المكمش للدجاج (1963-1948)

السنة	الاستهلاك (بلايين الدولارات)	السعر المكمش للرطل (بالسنت)
١٩٤٨	١٨,٣	٧٥,٤
١٩٤٩	١٩,٦	٧٥,٤
١٩٥٠	٢٠,٦	٧١,٨
١٩٥١	٢١,٧	٦٨,٠
١٩٥٢	٢٢,١	٦٦,٠
١٩٥٣	٢١,٩	٦٥,٠
١٩٥٤	٢٢,٨	٥٦,٤
١٩٥٥	٢١,٣	٥٨,٧
١٩٥٦	٢٤,٤	٥٠,٤
١٩٥٧	٢٥,٥	٤٧,٦
١٩٥٨	٢٨,٢	٤٥,٨
١٩٥٩	٢٨,٩	٤١,٤
١٩٦٠	٢٨,٢	٤١,٤
١٩٦١	٣٠,٣	٣٧,٠
١٩٦٢	٣٠,٢	٣٨,٦
١٩٦٣	٣٠,٦	٣٧,٦

كُمّش السعر باستخدام الرقم القياسي لأسعار المستهلكين ١٩٥٧ - ١٩٥٩=١٠٠

المصدر: Frederick V. Waugh. *Demand and Price Analysis-Some Examples from Agriculture* (Washington, D.C : U.S. Department of Agriculture, Thechnical Bulletin 1316, Nov. 1964), Table 5-1, p. 39

وبتطبيق هذا المنهج على المعادلة (3.106) على البيانات الموجودة في الجدول (٣-٨) نحصل على:

$$\begin{aligned} \ln(Q_t) &= \underset{(01.3)}{5.87} - \underset{(0.03)}{0.68}\ln(P_t) = \\ R^2 &= 0.97 \end{aligned} \tag{3.107}$$

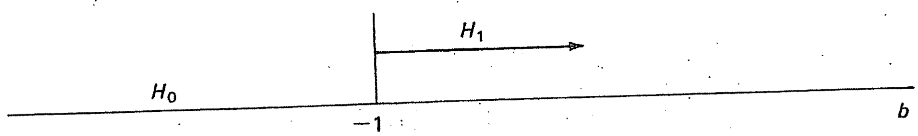
حيث الأرقام داخل الأقواس تحت المعاملات المقدرة هي الأخطاء المعيارية المقدرة المناظرة. وبهذا، يكون تقدير مرونة النقطة للطلب على الدواجن 0.68– وتقديرنا للمعلمة a بدرجة دقة لثلاثة أرقام هو ($e^{5.57} = 354$).

والآن، نوضح بعض الطرق التي عرضناها خلال هذا الفصل بدلالة نتائج المعادلة (3.107). افترض، على سبيل المثال، أننا نهتم باختبار الفرضية بأن الطلب على الدجاج غير مرتبط بسعره عند مستوى معنوية 5% مقابل الفرضية البديلة بأن الطلب على الدجاج يتأثر بالسعر. حينئذ تكون فرضية العدم $H_0 : b=0$ والفرضية البديلة$H_1 : b\neq0$، وبأخذ القيمة المطلقة لتقدير b في المعادلة (3.107) وبقسمته على الخطأ المعياري المقدر المناظر (0.68/0.03)، ينتج عنه نسبة عالية جداً لـ t تزيد على 20. وباستخدام قاعدتنا التجريبية للحساب يمكننا أن نستنتج مباشرة أن الفرضية $H_0 : b=0$ سترفض وفقا للنتائج في المعادلة (3.107).

وهناك توضيح آخر، افترض أننا نهتم باختبار فرضية أن الطلب على الدجاج مرن سعرياً عند واحد في المائة مستوى معنوية، مقابل الفرضية البديلة بأنه غير مرن. من مبادئ الاقتصاد الجزئي، نعرف أن الطلب يكون مرنا سعريا إذا كان معامل المرونة أقل من (1-) أو أكبر من واحد (بالقيمة المطلقة). وبدلالة نموذج الطلب بالمعادلة (3.105)، يناظر هذا أي قيمة لـ b (بحيث تكون $b<-1$). وهكذا فإن فرضيتنا للعدم والفرضية المقابلة في هذه الحالة سيكونان $H_0 : b < -1$ و $H_1 : b > -1$.

وبخلاف التوضيحات المعطاة في الأجزاء السابقة لهذا الفصل، فإن فرضيتنا للعدم لاتحدد قيمة معينة لـ b. وعلى الرغم من ذلك، فإن هذه الفرضية يمكن اختبارها بوساطة منهج فترة الثقة. ولنرى ذلك، لاحظ أن H_1 هي اختبار الذيل

الواحد، الذي يفيد أن قيمة b أكبر من جميع القيم الممكنة لها والمحددة بوساطة H_0. ويصور الشكل (٣-٨) هذه الحالة. وانطلاقا من نقاشنا السابق في هذا الفصل نبني فترة ثقة بذيل واحد وبحد أدنى، مثلا LB. وستأخذ هذه الفترة شكل $b \geq LB$. إذا كانت $LB \geq -1$ فإننا نرفض H_0، وإذا كانت $LB < -1$ فإننا نقبل H_0. لاحظ أن منهج الاختبار هذا يؤدي آليا إلى قبول H_0، وما رأيناه هو نشوء صعوبة إذا كانت فرضيتنا للعدم تأخذ الشكل $(H_0 : b \neq 0)$.



شكل (٣-٨)

وطريقة التحليل هذه مباشرة وواضحة. فباستخدام الجدول الإحصائي رقم٢ لتوزيع t بدرجات حرية قدرها (16-2=14)، نجد أن فترة الثقة 99% ذات الذيل الواحد هي:

$$b > \left(\hat{b} - t_{14;0.99}\hat{\sigma}_{\hat{b}}\right) = -0.68 - (2.624)(0.03) = -0.76$$

وطالما أن $-1 < -0.76$، فإننا نرفض H_0 ونقبل الفرضية البديلة بأن الطلب على الدجاج ليس مرنا سعريا.

اعتبر مرة أخرى النتائج في المعادلة (3.107)، ولاحظ أن المعادلة المقدرة «توفق» البيانات الملاحظة دائما. فالمعادلة تفسر 97% من التغير المشاهد في الاستهلاك الفردي السنوي المتوسط من الدجاج. ولكنك قد لاتشعر بالارتياح، وينبغي لك، في الأقل، لسببين، الأول: تفسر معادلتنا التغيرات السنوية في استهلاك الدجاج الناجمة عن تغيرات السعر، ولكننا نعلم أنه على مدى الفترة من ١٩٤٨ إلى ١٩٦٣م زادت دخول المستهلكين، فإذا كان الدجاج سلعة عادية، فإننا نتوقع أن

ارتفاع الدخول ينبغي أن ينسب إليه بعض الزيادة في استهلاك الدجاج خلال تلك الفترة. ولكن المعادلة (3.105) تهمل أي إشارة إلى تأثير ارتفاع الدخل. وفي هذه الحال، فإن حذف متغير مهم (الدخل) ينسب إلى تأثير السعر ليس، فقط، على استهلاك الدجاج ولكن، أيضا، على الدخل. لهذا السبب فإن قيمتنا المقدرة تبالغ في تأثير متغير السعر على المشتريات الاستهلاكية من الدجاج. والثاني: أننا، في منهجنا للتقدير، لم نأخذ في الحسبان جانب العرض من السوق، حيث إن الأسعار المشاهدة والكميات ليس نتاجا لتأثير الطلب، فقط، ولكنها أيضا، نتاج تفاعل الطلب والعرض. ينبغي أن ندخل ذلك صراحة في نموذج الانحدار، وهذا ماسنتناوله في الفصول المتبقية.

أسئلة

١- يُدّعي الآن أن الأداء في اختبارات I.Q قد تحسن في السنوات الأخيرة، وأن الوسط الحسابي للأداء الآن فوق ١٠٠. افترض أن الدرجات المتحصل عليها في I.Q موزعة توزيعا طبيعيا عبر المجتمع الإحصائي. فإذا كانت لدينا عينة مكونة من ١٠٠ اختبار حيث القيمة المتوسطة ١١٠والتباين المقدر هو ٤. اختبر فرضية أن القيمة المتوسطة للمجتمع الإحصائي I.Q أكبر من ١٠٠ عند مستوى معنوية ٥٪.

٢- اشرح لماذا يكون الوسط الحسابي للمجتمع المحسوب على أساس القيمة المتوسطة لعينة عشوائية مكونة من ٣٠ مشاهدة أفضل من تلك العينة المكونة من ٢٠ مشاهدة. هل كلا المقدرين غير متحيزين.

٣- افترض أن أسبوع العمل المعتاد في الصناعة هو ٤٠ ساعة. وفرضيتنا هي أنه، إذا اختلفت الساعات عن ٤٠ فسوف تميل للعودة إلى ٤٠ مرة أخرى. احدى طرق تكوين ذلك هو $\Delta H_t = B + \alpha(40 - H_{t-1}) + u_t$ حيث إن $\Delta H_t = H_t - H_{t-1}$، وهكذا فإن $H_t = (B + 40\alpha) + (1 - \alpha)H_{t-1} + u_t$. افترض أننا نقدر الانحدار التالي من البيانات ربع السنوية للولايات المتحدة الأمريكية:

$$\hat{H}_t = \underset{(0.7)}{5} + \underset{(0.15)}{0.875} H_{t-1}, \quad R^2 = 0.98$$

حيث تظهر الانحرافات المعيارية المقدرة داخل الأقواس تحت المعاملات المقدرة. هل ينبغي أن نقبل أو نرفض الفرضية بأن التغير في الساعات يعتمد على الانحراف عن ٤٠؟ لماذا ؟

٤- وضع السيد (أ) نظرية تقول إن الوسط الحسابي لأطوال الأفراد هو ٧٠ بوصة بينما يجادل السيد (ب) بأن نظرية السيد (أ) تبالغ في تأثير عوامل معينة، ومن ثم، فهي تبالغ في القيمة المتوسطة. افترض أن الأرقام التالية هي نتائج عينة عشوائية ذات حجم ٤ (٦٢ ، ٧٢ ، ٧٤ و ٦٤). افترض أن دالة الكثافة الاحتمالية موضع الاعتبار طبيعية وذات $\sigma^2 = 4$. اختبر نظرية السيد (أ) عند مستوى معنوية ٥٪.

٥- ما أهمية افتراضنا بأن الخطأ العشوائي موزع توزيعاً طبيعياً ؟

٦- افترض أن الوسط الحسابي لأطوال الأفراد في الجانب الشرقي من الولايات المتحدة الأمريكية هو ٦٧ بوصة. افترض أن صانعا للملابس الجاهزة في الشرق يريد أن يفتح متجرا لبيع الملابس في الجانب الغربي من الولايات المتحدة الأمريكية، وهو يعتقد أن الأفراد في الغرب أطول من الأفراد في الشرق. وإذا كان الأمر كذلك فإن عليه أن يصنع ملابس جاهزة أطول بعض الشئ. افترض أن هذا يتطلب إعادة تجهيز مكلف للآلات. وافترض أيضاً أن هذه الفرضية المرتبطة بالطول النسبي قد اختيرت بدلالة عينة عشوائية من أفراد إختيروا من الغرب. ناقش نتائج الخطأ من النوع الأول والخطأ من النوع الثاني.

٧- إشرح لماذا لايكون نموذج الانحدار الخطي مقيدا تقييدا كبيرا، على الرغم من أن كثيراً من العلاقات الاقتصادية غير خطية. اختر تحويلا ملائما ثم اشتق مصفوفة المشاهدات للنموذج:

$$Y_t = a + b\left(\frac{1}{1 - X_t}\right) + u_t$$

عندما تكون $n=3$، ومشاهداتنا عن Y و X هي$Y_1 = 1$، $Y_2 = 10$، $Y_3 = 12$
و $X_1=0$ ، $X_2=0.1$ و $X_3=0.5$.

٨- افترض أن أحد الأفراد قدر معادلة للاستهلاك وظهرت النتائج على النحو:

$$\hat{C} = \underset{(3.1)}{15} + \underset{(18.7)}{0.81} Y_d \quad \underset{R^2=0.99}{n=19}$$

حيث الأرقام الموجودة داخل الأقواس هي نسب t.

(ا) استخدم نسبة t لاختبار الفرضية بأن Y_d هو متغير مهم معنويا.

(ب) حدد الانحرافات المعيارية المقدرة لمقدرات المعلمات.

(جـ) كون فترة ثقة 95% لمعامل Y_d. هل تشتمل هذه الفترة على الصفر؟

٩- اعتبر الموقف الذي تؤدي فيه الزيادة في معدل المنافع للضمان الاجتماعي (مقاسة بالمدفوعات لكل أسرة في الشهر) إلى زيادة في الطلب على الضمان الاجتماعي عن طريق إحلال الأفراد للفراغ محل العمل. افترض، أيضا، أن الزيادة في الطلب على الضمان الاجتماعي تؤدي، بدورها، من خلال الضغوط السياسية، للزيادة في معدل المنافع للضمان الاجتماعي في الفترة التالية. عبر عن هذه العلاقات باستخدام نموذج مكون من معادلتين.

١٠- اعتبر الحال التي يعتمد فيها عدد المنشآت الاقتصادية التي تستوطن ولاية معينة على معدل الضريبة النسبي في هذه الولاية. افترض، أيضا، أنه، على الرغم من وجود منافع ضريبية للمقيمين في الولاية، فإن زيادة عدد المنشآت التي تستوطن ولاية معينة تؤدي إلى زيادة معدل التلوث بها. عبر عن العلاقات بين مكان توطن المنشأة والتلوث بدلالة نماذج الانحدار.

١١- اعتبر نموذج الانحدار المعتاد التالي:

$$Y_t = a + bX_t + u_t$$

افترض أنه لايمكننا قياسX_t، وافترض، بدلا من ذلك أننا نشاهد المتغير Z_t حيث $Z_t=5\text{-}3X$، ما معلمات النموذج التي تربط بين Y_t و Z_t ؟

الفصل الرابع

تحليل الانحدار المتعدد

استكشفنا في الفصول السابقة العلاقة الإحصائية بين متغيرين، ففي حالة الإنفاق الاستهلاكي (مثلا) استخدمنا النموذج:

$$C_t = a + bY_{dt} + u_t, \tag{4.1}$$

وطورنا الطرق التي يمكننا عن طريقها تقدير كل من a و b، واختبار الفرضيات حول هذه العلاقة. إلا أننا، عادة، نجد أن العلاقات الاقتصادية أكثر تعقيدا من هذا النموذج (4.1) بمعنى أن قيمة متغير معين كالاستهلاك تعتمد ليس، فقط، على متغير واحد، وانما على مجموعة كاملة من المتغيرات المستقلة.

افترض، على سبيل المثال، أن مستوى الإنفاق الاستهلاكي في الفترة t يعتمد ليس فقط، على مستوى الدخل المتاح الحالي، ولكن، أيضا، على قيمة الأصول السائلة (A_t) * وعلى الدخل المتاح في الفترة السابقة $Y_{d(t-1)}$. فإذا كان مايمتلكه المستهلك من الأصول السائلة كبير جدا غير عادي (كما كان الحال عندما شارفت الحرب العالمية الثانية على الانتهاء)، فسوف نتوقع أن يكون الإنفاق الاستهلاكي أكبر، نوعا ما، عن ذلك المرتبط عادة بمستوى الدخل السائد. وعلى العكس إذا كان المخزون من الأصول السائلة منخفضا انخفاضا غير عادي، فإن

---

* نعني بالأصول السائلة مايملكه الفرد من النقود والودائع الآجلة والمدخرات، وانصبة القروض، والسندات الحكومية.

المستهلكين قد يقللون إنفاقهم الاستهلاكي بعض الشئ من أجل استكمال النقص في ممتلكاتهم من الأصول السائلة. وقد يؤدي الدخل السابق دورا في تحديد المستويات الحالية للإنفاق الاستهلاكي، فمثلا قد تكون بعض النفقات التي تتم في الفترة الحالية محفزة بوساطة المستويات السابقة للمعيشة. كما قد ترتبط بمستوى الدخل في الفترة السابقة مباشرة. فعلى سبيل المثال، مع بقاء العوامل الأخرى على حالها، إذا كان الدخل أعلى في الفترة السابقة مباشرة، فإن من المرجح أن يرتفع مستوى الاستهلاك الحالي. وكل هذا يعني أنه لدينا الآن دالة استهلاك معقدة مثل:

$$C_t = a + b_1 Y_{dt} + b_2 A_t + b_3 Y_{d(t-1)} + u_t. \qquad (4.2)$$

ومشكلتنا الآن هي تقدير كيفية اعتماد الاستهلاك على جميع هذه المتغيرات المستقلة، وبمعنى آخر، ينبغي أن توجد طريقة يمكن بوساطتها تقدير قيم a_1، b_1، b_2 و b_3.

ومرة أخرى، نود معرفة تباين مقدراتنا حتى نستطيع الحصول على مقياس لدقة هذه المقدرات. وعلى سبيل تعميم مباشر لما عرفناه من قبل، سنفترض أن لدينا مجموعة من القيم المشاهدة لمتغيرات نموذجنا، فعلى سبيل المثال، وبالمقابل مع المعادلة (4.2) فإن لدينا معلومات تقابلها في الجدول رقم (٤-١).

للوهلة الأولى، تظهر طبيعة مشكلتنا مختلفة بعض الشئ وأكثر تعقيدا مقارنة بحالة الانحدار البسيط. وبالتحديد، ففي حالة الانحدار البسيط، لدينا قيم مشاهدة عن المتغير التابع ومتغير مستقل واحد، وكانت مشكلتنا، ببساطة، هي تقدير كيف يتغير الأول مع المتغير الثاني، والآن، لدينا مشاهدات عن مجموعة كاملة من المتغيرات المستقلة، ونواجه مهمة أكثر صعوبة تتمثل في معرفة آثار المتغيرات المستقلة كافة. أي أنه لتحديد، تأثير كل متغير مستقل، ينبغي علينا أن نعزل، بدرجة ما، تأثيره على المتغير التابع عن تأثير باقي المتغيرات المستقلة الأخرى.

جدول رقم (١ , ٤) القيم المشاهدة (ببلايين الدولارات)

السنة	الاستهلاك (C_t)	الدخل المتاح (Y_{dt})	الأصول المالية (A)	الدخل المتاح من السنة السابقة $Y_d(t-1)$
١٩٦٠	٥٢٥,٢	٣٥٠,٠	٣٩٩,٢	٣٣٧,٣
١٩٦١	٣٣٥,٢	٣٦٤,٤	٤٢٤,٦	٣٥٠,٠
١٩٦٢	٣٥٥,١	٣٨٥,٣	٤٩٩,٠	٣٦٤,٤
١٩٦٣	٣٧٥,٠	٤٠٤,٦	٤٩٥,٤	٣٨٥,٣
١٩٦٤	٤٠١,٢	٤٣٨,١	٥٣٠,٥	٤٠٤,٦
١٩٦٥	٤٣٢,٨	٤٧٣,٢	٥٧٣,١	٤٣٨,١
١٩٦٦	٤٦٦,٣	٥١١,٩	٦٠١,٥	٤٧٣,٢
١٩٦٧	٤٩٢,١	٥٤٦,٣	٦٥٠,٤	٥١١,٩
١٩٦٨	٥٣٥,٨	٥٩١,٢	٧٠٩,٦	٥٤٦,٣
١٩٦٩	٥٧٧,٥	٦٣١,٦	٧٣١,٦	٥٩١,٢

المصدر: التقرير الاقتصادي للرئيس (واشنطن، مكتب الطباعة الحكومية في الولايات المتحدة الامريكيـة، فبراير ١٩٧١م) الصفحات ١٩٧ ، ٢٠٤ و ٢٦٢.

وعلى الرغم من هذا التعقيد الظاهري، فإننا نؤكد في البداية أن تحليل الانحدار المتعدد (أي الحال التي يكون لدينا فيها أكثر من متغير مستقل واحد) هو تعميم مباشر للتحليل الثنائي المتغيرات. سنعرض في البداية نموذج انحدار متعدد يحقق الافتراضات المرتبطة بخصائص الخطأ العشوائي كافة التي كوناها في حال انحدار المتغيرين. بعد ذلك، سوف نتبع طريقة المتغير المساعد التي، عن طريقها نضع افتراضاتنا بوصفها شروطا على مقدر الخطأ العشوائي. سينتج عن ذلك مجموعة من المعادلات الطبيعية، وبحل هذه المعادلات، نحصل على مقدرات غير متحيزة لكل المعاملات في نموذجنا. وبهذا، يكون منهجنا تقريبا هو المنهج نفسه المستخدم في الفصول السابقة. فإذا كنت قد فهمت حالة الانحدار البسيط فلن تواجه صعوبة كبيرة في فهم تحليل الانحدار المتعدد.

(٤-١) نموذج الانحدار المتعدد

عموما، دعنا نعتبر نموذج الانحدار التالي:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + \cdots + b_k X_{kt} + u_t, \quad t=1,\cdots,n, \tag{4.3}$$

حيث:

Y_t = المشاهدة رقم t عن المتغير التابع.

X_{it} = المشاهدة رقم t عن المتغير المستقل X_i، حيث $(i=1,2,\dots,k)$ إذا كان لدينا مشاهدات مستقلة عددها k.

u_t = القيمة t للخطأ العشوائي، وأخيرا.

b_i = معامل المتغير المستقل رقم i.

سوف يكون من الملائم بدءا من الآن استخدام b_0 (بدلا من a) لنرمز إلى الحد الثابت في معادلة الانحدار المتعدد حتى تأخذ المعلمات كافة في المعادلة (4.3) الشكل b's. سنعرض الآن قائمة بالافتراضات التي نكونها لنموذج الانحدار المتعدد. وتتطلب غالبية هذه الافتراضات تعليقا محدودا طالما أنها الافتراضات نفسها التي فرضناها في حالة الانحدار البسيط:

١- القيمة المتوقعة أو المتوسطة للخطأ العشوائي هي الصفر

$$E(u_t)=0.$$

٢ - تباين الخطأ العشوائي ثابت، ولذا يكون مستقلا عن

$$E(u_t-0)^2 = E(u_t)^2 = \sigma_u^2.$$

٣ - قيم الخطأ العشوائي مستقلة عن بعضها بعضا، لذا، يكون التغاير بين الأخطاء العشوائية المناظرة لأي مشاهدتين (u_t, u_s) صفرا،

$$\text{cov}\,(u_t, u_s) = E(u_t u_s) = 0.$$

٤ - استقلال الخطأ العشوائي عن جميع قيم المتغيرات المستقلة. وبالتحديد نفترض أن u_t مستقل عن $X_{1s}, \dots, X_{ks}$ لجميع s, t. وينتج عن ذلك أن التغاير بين الخطأ العشوائي u_t وكل واحد من المتغيرات المستقلة لمعادلتنا للانحدار (4.3) هو الصفر. ومرة ثانية، فإن هذا يعني أنه، سواء وضعت قيم المتغيرات

المستقلة بوساطة «الباحث» الذي يجري التجارب أو بوساطة «الاقتصاد» فإن تلك القيم لاتؤثر بأي طريقة كانت على قيم الخطأ العشوائي. واصطلاحا يمكن التعبير عن شرط التباين المشترك بين u_t وكل من X_{it} على النحو:

$$\operatorname{cov}(u_t, X_{it}) = E\left[u_t\left(X_{it} - \mu_{Xi}\right)\right] = E(u_t X_{it}) - \mu_{Xi}E(u_t) = E(u_t X_{it}) = 0$$

حيث

$$E(X_{it}) = \mu_{Xi}$$

٥ - لايوجد ارتباط خطي تام بين المتغيرات المستقلة. بمعنى أن أي المتغيرات المستقلة ليس توليفة خطية من المتغيرات الأخرى. على سبيل المثال، تستبعد هذه القاعدة علاقات مثل:

$$\text{a. } X_{1t} = 3 - 2X_{2t} + 17X_{3t};$$

$$\text{b. } X_{4t} = (X_{1t} + X_{2t} + X_{3t})/3; \text{ or}$$

$$\text{c. } X_{2t} = 3X_{8t}.$$

ولكننا لانستبعد العلاقات غير الخطية، فمثلا إذا كانت $X_{1t} = Y_{2t}^2$، وإذا كانت $X_{3t} = X_{5t}\, X_{6t}$، فإن افتراضنا سيظل صحيحا.

تعرفت من قبل على الافتراضات الأربعة الأولى من هذه الافتراضات من نموذج الانحدار البسيط، وتبريرها هو التبرير نفسه الذي اوردناه في الحالة السابقة. فإذا لم تكن متأكدا من سبب تكوين هذه الافتراضات الأربعة فعليك أن تجدد ذاكرتك بالعودة إلى المناقشة الموجودة في حالة الانحدار البسيط في الفصل الثاني.

أما الافتراض الجديد الذي أضفناه هنا فهو الافتراض الخامس . وهو، في الحقيقة، امتداد للافتراض الذي كوناه في الفصل الثاني بأن المتغير المستقل X_t في حالة انحدار المتغيرين ينبغي أن يكون له، في الأقل، قيمتان مختلفتان. رأينا في الفصل الثاني في حال انحدار المتغيرين إذا كانت قيمة X لاتتغير، أي أن $X_t \equiv X_0$، فإننا فقط نستطيع تقدير معلمة واحدة، والتي نطلق عليها A، تعتمد على كل من الحد الثابت الأصلي، a، وعلى ماهو، في الحقيقة، حد ثابت أوجد بوساطة

القيمة غير المتغيرة لـــX_t وbX_0، أي $A = a + bX_0$. وباختصار، إذا كانت قيمــة المتغير المستقل لاتتغير مطلقا، فإن تأثيره على Y لايمكن فصله عن تأثير الحد الثابت الأصلي.

في حال الانحدار المتعدد هذه، نرغب في ضمان ليس، فقط، أنb_i أو أن تأثيرX_{it} على Y، يمكن عزله عن الحد الثابتb_0، ولكن أيضا يمكن عزله عن تأثير المتغيرات المستقلة الأخرى كافة. وتظهر قوة الافتراض الخامس في أنه يضمن مثل هذا العزل. وسنبين هذا بدقة في المبحث (٤-٢)، ولكننا نرغب الآن في توضيح ماذا يمكن أن يترتب على انتهاك الافتراض الخامس.

افترض أنX_{1t} تأخذ قيما متغيرة ولكنها تتعادل دائما مع X_{2t}. حينئذ (ومع بقاء العوامل الأخرى على حالها)، إذا تزايدX_{1t} بمقدار وحدة واحدة فإنY_t سوف يتغير بمقدار$(b_1 + b_2)$ وحدة لأنه إذا كانX_{it} يتعادل دائما معX_{2t}، فإنX_{2t} يتزايد أيضا، بمقدار وحدة واحدة ! يوحي هذا بأن التأثير المؤلــف لـــX_{1t} وX_{2t} (وهــو $(b_1 + b_2)$ هو الذي يمكن تقديره إذا كانت قيم $X_{1t} = X_{2t}$. وببساطة لاتوجد طريقة يمكن بوساطتها عزل تأثيــر X_1 من تأثيــر X_2 على Y. ويتبع هذا بسبب أنــه، إذا كانت$X_{1t} = X_{2t}$، فإن نموذجنا الأساسي بالمعادلــة (4.3) يمكن أن تعاد كتابته على النحو:

$$Y_t = b_0 + BX_{1t} + b_3X_{3t} + \cdots + b_kX_{kt} + u_t, \qquad (4.4)$$

حيث إن $B = (b_1 + b_2)$. إلا أنه، إذا كان المتغيران المستقلان متساويين دائما مــع بعضهما بعضا فإنه يمكن حذف واحد من المتغيرين من نموذجنا دون خسارة فــي المعلومات. وحينئذ، يمكن اعتماد النموذج الناتج (أو المختزل) والذي يحتوي على (k-1) متغيرات مستقلة، وسيحتوي هذا النموذج على المعلمــة B السابق الإشارة إليها، التي تعبر عن الأثر المؤلف للمتغيرين المستقلين الأصلين في المعادلة. لاحظ أن المعادلة (4.4) لاتقترح عدم قابلية b_0، b_3، ... ، b_k للتقدير ومرة أخرى سنعود إلى هذه المشكلة بدقة أكثر في المبحث (٤-٢).

وهكذا، فإن نموذج الانحدار المتعدد مشابه جدا لنموذج الانحدار البسيط، فهو يصف العلاقة الدالية الخطية والتي تتحدد عن طريقها قيمة المتغير التابع بوساطة قيم مجموعة من المتغيرات المستقلة. والآن ينبغي علينا أن نوجد طريقة لتقدير قيم المعلمات في هذه العلاقة.

(٤-٢) التقدير بوساطة المتغيرات المساعدة

تذكر أن طريقة المتغيرات المساعدة تتضمن فرض مجموعة من الشروط على مقدر الخطأ العشوائي التي تقترحها افتراضات نموذج الانحدار. ففي حالة الانحدار البسيط فرضنا الشرطين التاليين:

الافتراض المناظر	الشرط
$E(u_t) = 0$	1. $\Sigma \frac{\hat{u}_t}{n} = 0,$ أو $\Sigma \hat{u}_t = 0$
$E(X_t u_t) = 0$	2. $\Sigma \frac{\left(X_t \hat{u}_t\right)}{n} = 0,$ أو $\Sigma\left(X_t \hat{u}_t\right) = 0$

وبمساعدة هذين الشرطين، أوجدنا معادلتين طبيعيتين، حللناهما للحصول على قيم $\hat{a}$ و $\hat{b}$. وهنا سوف نستخدم المنهج نفسه ولكن قبل أن نقوم بذلك، ينبغي علينا أن نستخدم بعض تعريفاتنا الأساسية السابقة ضمن اطار الانحدار المتعدد. فعلى سبيل المثال، يعتمد منهجنا في التقدير اعتمادا حاسما على مقدر الخطأ العشوائي $\hat{u}_t$. لذلك ينبغي علينا تعريف $\hat{u}_t$ ضمن إطار نموذج الانحدار المتعدد.

إذا كان نموذجنا للانحدار هو (4.3)، فإن القيمة المتوسطة لـ Y_t المرتبطة بقيم X_{1t}، ...، X_{kt} هي:

$$Y_t^m = b_0 + b_1 X_{1t} + \cdots + b_k X_{kt}. \tag{4.5}$$

وكما هو الحال في حالة الانحدار البسيط يمكن كتابة Y بوصفها مجموع كل من قيمتها المتوسطة والخطأ العشوائي:

$$Y_t = Y_t^m + u_t. \tag{4.6}$$

وإذا عرفنا b_0 ، b_1، ...، b_k،، يمكننا أن نشتق قيمة u_t على النحو:

$$u_t = Y_t - Y_t^m. \tag{4.7}$$

لاحظ أن المعادلتين (4.6) و (4.7) تتماثلان مع نظيراتهما في حال الانحدار البسيط.

افترض أن لدينا مقدرات لـ b_0 ، b_1 ، . . . ، b_k مثلا $\hat{b}_0$ ، $\hat{b}_1$ ، . . . ، $\hat{b}_k$ في ضوء المعادلة (4.5)، يصبح مقدرنا للقيمة المتوسطة لـY_t:

$$\hat{Y}_t = \hat{b}_0 + \hat{b}_1 X_{1t} + \cdots + \hat{b}_k X_{kt}, \tag{4.8}$$

لقد أسقطنا، أيضا في هذه المرة الرمز العلوي m من أجل التبسيط. وحينئذ يكون مقدرنا للخطأ العشوائي المقترح بوساطة المعادلة (4.7) هو:

$$\hat{u}_t = Y_t - \hat{Y}_t \tag{4.9}$$

$$= Y_t - \hat{b}_0 - \hat{b}_1 X_{1t} - \cdots - \hat{b}_k X_{kt}.$$

يمكننا إعادة كتابة المعادلة (4.9) على النحو:

$$Y_t = \hat{Y}_t + \hat{u}_t, \tag{4.10}$$

أو بطريقة متكاملة على النحو:

$$Y_t = \hat{b}_0 + \hat{b}_1 X_{1t} + \ldots + \hat{b}_k X_{kt} + \hat{u}_t. \tag{4.11}$$

باختصار، تناظر تعريفاتنا لكل من $\hat{Y}_t$ و$\hat{u}_t$ بدقة المقدرين أنفسهما في حالة نموذج الانحدار البسيط في الفصل الثاني. والآن، طالما يتوافر لنا تعريفات كافية فسنتجه إلى التقدير.

المعادلات الطبيعية

لإيجاد المعادلات الطبيعية في حالة نموذج الانحدار المتعدد، سنتبع المنهج نفسه الذي اتبعناه في إيجاد تلك المعادلات في حالة نموذج انحدار المتغيرين. والآن، وبافتراض وجود عدد k من المتغيرات المستقلة، سيصبح لدينا (k+1) شروط نفرضها على مقدر الخطأ العشوائي وبالتحديد، لدينا:

$$1.\ \sum\frac{\hat{u}_t}{n}=0,\ \text{أو}\ \sum\hat{u}_t=0 \qquad E(u_t)=0$$

$$2.\ \sum\frac{(X_{1t}\hat{u}_t)}{n}=0,\ \text{أو}\ \sum(X_{1t}\hat{u}_t)=0 \qquad E(X_{1t}u_t)=0$$

$$3.\ \sum\frac{(X_{2t}\hat{u}_t)}{n}=0,\ \text{أو}\ \sum(X_{2t}\hat{u}_t)=0 \qquad E(X_{2t}u_t)=0$$

$$\vdots$$

$$k+1.\sum\frac{(X_{kt}\hat{u}_t)}{n}=0,\ \text{أو}\ \sum(X_{kt}\hat{u}_t)=0 \qquad E(X_{kt}u_t)=0$$

لاحظ أن لدينا الآن (k+1) شروط و (k+1) معلمات هي b_0 , b_1 , ..., b_k التي تظهر في نموذجنا للانحدار.

دعنا الآن نفحص كيف تولد هذه الشروط عدد (k+1) من المعادلات الطبيعية.

إذا جمعنا أولا معادلة (4.1) على مدى جميع المجموعات التي عددها n للقيم المشاهدة، يكون لدينا:

$$\sum Y_t=n\hat{b}_0+\hat{b}_1\sum X_{1t}+\hat{b}_2\sum X_{2t}+\cdots+\hat{b}_k\sum X_{kt}+\sum\hat{u}_t. \qquad (4.12)$$

وبفرض الشرط $\Sigma\hat{u}_t=0$ ، نحصل على معادلتنا الطبيعية الأولى.

$$\text{N1.}\quad \sum Y_t=n\hat{b}_0+\hat{b}_1\sum X_{1t}+\hat{b}_2\sum X_{2t}+\cdots+\hat{b}_k\sum X_{kt}.$$

نحصل على k معادلات طبيعية إضافية عن طريق ضرب المعادلة (4.11) بكل من المتغيرات المستقلة التي عددها k والجمع على مدى المجموعات التي عددها n من القيم المشاهدة. وبالتحديد،

$$\sum(X_{1t}Y_t)=\hat{b}_0\sum X_{1t}+\hat{b}_1\sum X_{1t}^2+\cdots+\hat{b}_k\sum(X_{1t}X_{kt})+\sum(X_{1t}\hat{u}_t),$$

$$\sum(X_{2t}Y_t)=\hat{b}_0\sum X_{2t}+\hat{b}_1\sum(X_{2t}X_{1t})+\cdots+\hat{b}_k\sum(X_{2t}X_{kt})+\sum(X_{2t}\hat{u}_t),$$

$$\vdots$$

$$\sum(X_{kt}Y_t)=\hat{b}_0\sum X_{kt}+\hat{b}_1\sum X_{kt}X_{1t})+\cdots+\hat{b}_k\sum X_{k_t^2}+\sum(X_{kt}\hat{u}_t).$$

إذا فرضنا الشروط $\Sigma(X_{it}\hat{u}_t)=0$ لكل $(i=1,2,...,k)$، يسقط الحد الأخير من كل من هذه المعادلات، ونحصل، بذلك، على المعادلات الطبيعية الباقية (التي يبلغ عددها k).

$$\text{N2.}\qquad \sum(X_{1t}Y_t)=\hat{b}_0\sum X_{1t}+\hat{b}_1\sum X_{1t}^{\ 2}+\cdots+\hat{b}_k\sum(X_{1t}X_{kt})$$

$$\text{N2.}\qquad \sum(X_{1t}Y_t)=\hat{b}_0\sum X_{1t}+\hat{b}_1\sum(X_{2t}X_{1t})+\cdots+\hat{b}_k\sum(X_{1t}X_{kt})$$

.

$$\text{N(k+1).}\sum(X_{kt}Y_t)=\hat{b}_0\sum X_{kt}+\hat{b}_1\sum(X_{kt}X_{1t})+\cdots+\hat{b}_k\sum X_{kt}$$

لاحظ أن عمليات الجمع في المعادلات الطبيعية أعلاه تعتمد، فقط، على قيم المتغير التابع والمتغيرات المستقلة، فقط. وهكذا فحالما تتوافر لدينا عينة من المشاهدات فإنه يمكن حساب قيم هذه المجاميع. ويترتب على ذلك، أنه يمكن النظر إلى المعادلات الطبيعية أعلاه على أنها مجموعة مكونة من (k+1) من المعادلات في (k+1) من المجاهيل $\hat{b}_0$، $\hat{b}_1$، ...، $\hat{b}_k$ وسنكون، بعامة، قادرين على تحديد قيم هذه المقدرات عن طريق حل المجموعة السابقة من المعادلات الطبيعية.

للتوضيح، اعتبر حالة الانحدار المتعدد الذي تحتوي على متغيرين مستقلين:

$$Y_t=b_0+b_1X_{1t}+b_2X_{2t}+u_t.$$

سينتج، باتباع الطرق الموصوفة أعلاه، مجموعة من ثلاث معادلات طبيعية:

$$\sum Y_t=n\hat{b}_0+\hat{b}_1\sum X_{1t}+\hat{b}_2\sum X_{2t},$$

$$\sum(X_{1t}Y_t)=\hat{b}_0\sum X_{1t}+b_1\sum X_{1t}^2+\hat{b}_2\sum(X_{1t}X_{2t}),$$

$$\sum(X_{2t}Y_t)=\hat{b}_0\sum X_{2t}+b_1\sum(X_{2t}X_{1t})+\hat{b}_2\sum X_{2t}^2.$$

افترض أن حساباتنا مع القيم المشاهدة Y_t، X_{1t} و X_{2t} تعطينا:

$n=10$	$\sum X_{1t}=2$	$\sum X_{2t}=2$
$\sum X_{1t}^{\ 2}=6$	$\sum(X_{1t}X_{2t})=1$	$\sum X_{2t}^{\ 2}=4$
$\sum Y_t=5$	$\sum(X_{1t}Y_t)=6$	$\sum(X_{2t}Y_t)=7$

وبإدخال هذه القيم المحسوبة في المعادلات الطبيعية، نحصل على:

$$5=10\hat{b}_0+2\hat{b}_1+2\hat{b}_2,$$
$$6=2\hat{b}_0+6\hat{b}_1+\hat{b}_2,$$
$$7=2\hat{b}_0+\hat{b}_1+4\hat{b}_2.$$

ويعطينا حل هذه المجموعة من المعادلات:

$$\hat{b}_0=0.045, \qquad \hat{b}_1=0.727, \qquad \hat{b}_2=1.545.$$

ولذا، تكون معادلتنا المقدرة للانحدار هي:

$$\hat{Y}=0.045+0.727X_1+1.545X_2.$$

ويعطينا هذا تقديرا للقيمة المتوسطة لـ Y المرتبطة مع أي مجموعة محددة من القيم لـ X_1 ، X_2.

وكما قد تخمن، فإن الحل الفعلي لمجموعة المعادلات الطبيعية لتحديد القيم المقدرة للمعاملات يمكن أن يتضمن عددا ضخما من الحسابات، حتى وإن كان عدد المتغيرات صغيرا. لذا، ولأغراض عملية، عادة مايتطلب استخدام تحليل الانحدار المتعدد استعمال الحاسوب. إلا أنه، من حيث المبدأ، من المهم أن تفهم كيف تحدد القيم المقدرة للمعلمات. وسيساعد هذا ليس، فقط، في تفسير النتائج على نحو صحيح، ولكن (وبمساعدة بعض الموضوعات التي سنعرضها فيما بعد) في اكتشاف الصعوبات.

مشكلة تعدد العلاقات الخطية التام

قلنا في الافتراض الخامس أعلاه أن منهجنا للتقدير يفشل إذا كان واحد (أو أكثر) من المتغيرات المستقلة مولفا خطيا من المتغيرات المستقلة الأخرى. ونحن الآن في وضع يسمح لنا بمعرفة السبب. افترض (على سبيل المثـال) أن X_k في المعادلة (4.3) يساوي:

$$X_{kt}=c_0+c_1X_{1t}+c_2X_{2t}+\cdots+c_{k-1}X_{(k-1)t}, \tag{4.13}$$

حيث c_0، c_1، ...، c_{k-1} ثوابت. وبالمناسبة، فإن بعض هذه الثوابت قد يكون صفريا. تذكر من المناقشة السابقة أننا أوجدنا عدد (k+1) من المعادلات الطبيعية، عن طريق ضرب المعادلة (4.11) بـX_{kt} ومن ثم الجمع، ووضع $\Sigma(X_{kt}\,\hat{u}_t)=0$، فإذا كان X_{kt} يساوي التعبير الموجود في المعادلة (4.13) فإنه يمكننا اشتقاق المعادلة الطبيعية رقم (k+1) طريق ضرب المعادلة الطبيعية الأولى بوساطة c_0، والثانية في c_1 وهلم جرا، ثم جمعها بعد ذلك. فمثلا يمكننا في ضوء المعادلة (4.13) التعبير عن الجانب الأيسر من المعادلة الطبيعية (k+1) على النحو التالي:

$$\sum(Y_tX_{kt})=c_0\sum Y_t+c_1\sum(Y_tX_{1t})+\cdots+c_{k-1}\sum(Y_tX_{(k-1)t}).$$

نستحث القارئ أن يقنع نفسه بهذه القاعدة من خلال العمل مع مثال لنموذج الانحدار المتعدد بثلاثة متغيرات مستقلة.

ولكن هذا يعني أن المعادلة الطبيعية (k+1) ليست معادلة مستقلة، إنها توليفة خطية من المعادلات الـ (k) الأولى. في هذه الحال، نرغب في اشتقاق مقدرات لعدد (k+1) معلمات: b_0، b_1، ...، b_k، ولكن، تتوافر لدينا، فقط، معادلات مستقلة عددها k. ونتيجة لذلك لايمكننا (عموما) حل هذه المعادلات للحصول على مقدرات وحيدة لهذه المعلمات. وبدقة أكثر تنحصر مشكلتنا في أنه طالما أن واحدا من متغيراتنا المستقلة عادة مايكون مجموعا موزونًا من قيم المتغيرات المستقلة الأخرى، فلن نكون قادرين على عزل تأثيره على المتغير التابع من تأثير المتغيرات الأخرى.

ولكننا، في هذه الحال، يمكننا تقدير الآثار المجمعة لهذه المتغيرات. فمثلا، إذا قمنا بإحلال المعادلة (4.13) في المعادلة (4.3) فإننا نحصل على:

$$\begin{aligned} Y_t &= (b_0+b_kc_0)+(b_1+b_kc_1)X_{1t}+\cdots \\ &\quad +(b_{k-1}+b_kc_{k-1})X_{(k-1)t}+u_t \\ &= d_0+d_1X_{1t}+\cdots+d_{k-1}X_{(k-1)t}+u_t, \end{aligned} \tag{4.14}$$

حيث إنه (عموما) تكـون $d_i = (b_i + b_k c_i)$. في المعادلـة (4.14)، لدينا معلمـات عددها k: d_0، d_1، ...، d_{k-1}، فإذا لم يكن هناك علاقات أخرى من النوع الموجود في المعادلـة (4.13) فإن لدينا عـدد k من المتغيرات المستقلة التي تحقق افتـراضـنـا الخامس. باختصار، يمكننا الآن اشتقاق k من المعادلات الطبيعية وحلها للوصول إلى مقدرات وحيدة لـ $\hat{d}_0$، $\hat{d}_1$، ...، $\hat{d}_{k-1}$.

وطالما أن $d_i = (b_i + b_k c_i)$، فإنه لايمكننا أن نعامل مقدرنا لـ d_i كمقدر b_i أي لن نقدر، عموما، على تقدير تأثير X_i على Y. هناك حالة استثنائية وحيدة وهي عندما تكون $c_i = 0$. فعلى سبيل المثال، إذا كانت $c_5 = 0$ فإن $d_5 = b_5$، حيث يمكننا معاملة مقدرنا لـ d_5 مقدر لقيمة b_5. وبدلالة المعادلة (4.13) يمكننا أن نجد أن قيمة معينة لـ c_i تساوي الصفر إذا كانت قيمة المتغير X_k لاتعتمد على قيمة X_t. وعموما، لايمكننا تقدير معامل متغير مستقل منحل degenerate، مثـل X_{kt}، الذي هو توليفة خطية للمتغيرات المستقلة الأخرى في معادلة الانحدار. على سبيل المثال، إذا كان:

$$X_{2t} = 3 - 17X_{1t} + 8X_{5t},$$

حينئذ، فإنه، وباستثـنــاء ($c_0 = 3$، $c_1 = -17$ و $c_3 = 8$) فإن جميع c_i's تسـاوي الصفر. ونتيجة لذلك، يمكن تقدير جميع الـ c_i's الأصلية (باستثناء b_0، b_1، b_2 و b_5). باختصار، نحث القاري على ألا يحفظ هذه النتيجة عن ظهر قلـب، ولكن نطلب منه تعلم الطريقة التي اشتققناها: إذا كان واحد (أو أكـثـر) مـن المتغيرات المستقلة مرتبطا خطيا ببعض المتغيرات الأخرى نستخدم، ببـسـاطـة، العلاقات الخطية للإحلال محل المتغيرات المنحلة وبعد ذلك نقدر معلمات النموذج المختزل للانحدار. ثم نحدد أخيرا معلمات نموذج الانحدار الأصلي التي سنقوم بتقديرها عن طريق مقارنة معلمات نموذج الانحدار الأصلي مع معلمات معادلة الانحدار المختزل.

(٤-٣) خصائص المقدرات واختبار الفرضيات

تفسير المقدرات*

كما أكدنا، من قبل، خلال هذا الفصل، أنه، لتقدير تأثير متغير معين مثل X_k على Y، ينبغي علينا أن نعزل، أو نأخذ في الحسبان، تأثير المتغيرات المستقلة الأخرى كافة. وحينئذ، فقط، يمكننا عزل تأثير X_k على Y. وعلى الرغم من أن ذلك لم يكن واضحا من المجموعة السابقة من معادلاتنا الطبيعية، إلا أن ذلك هو مايفعله منهجنا في التقدير بالضبط. وقد استخدمنا في ملحق هذا الفصل، والذي نشجع القارئ بقوة على تتبعه، منهجا بديلا لاشتقاق صيغ مقدراتنا، $\hat{b}_i$. وبهذه الطريقة يمكنك معرفة الكيفية التي تستطيع من خلالها مناهجنا للتقدير فصل sorting تأثير مختلف المتغيرات المستقلة. إضافة إلى ذلك، من السهل، مع وجود الصيغ الصريحة للمقدرات في الملحق، إثبات أن هذه المتغيرات غير متحيزة، كما في حالة انحدار المتغيرين.

إلا إننا، عند هذه النقطة، نقدم مناقشة مختصرة معتمدة على القريحة لما رأيناه في الملحق. افترض اننا نحاول، باستخدام المعادلة (4.3)، تقدير تأثير تغير وحدة من X_k على Y، أو أننا، نحاول تقدير b_k بافتراض ثبات العوامل الأخرى.

نعلم أن b_k لن تكون قابلة للتقدير إذا كانت X_k توليفة خطية من المتغيرات الأخرى. افترض أن X_k ليست توليفة خطية وهذا لايتضمن بالطبع أن X_k غير مرتبطة، مع المتغيرات المستقلة الأخرى. على سبيل المثال ففي دالة الاستهلاك قد تكون X_k الأصول السائلة و X_1 قد يكون الدخل المتاح. فإذا كان الأمر كذلك فإننا نتوقع، بالتأكيد، أن ترتبط X_k طرديا مع X_1، ولكن، كما هو الحال بالنسبة لطول الفرد ووزنه، هذين المتغيرين لاينبغي أن يرتبطا ببعضهما ارتباطا تاما.**

---

* هذا المبحث صعب بعض الشئ، ولذا، فإنه باستثناء فهم القاعدة (4.19)، يمكن تجاهل المعلومات الأخرى في هذا المبحث عند القراءة الأولى بدون فقدان تواصل المعلومات.

** إلى حد ما، لايحدد دخل الفرد الأصول السائلة التي يملكها تمامًا، على سبيل المثال، فقد يكون هناك فردين لهما الدخل نفسه ولكن أصولهما السائلة مختلفة.

وإذا قمنا بتعيم النتيجة السابقة بعض الشئ نذكر في نموذجنا للانحدار (4.3) قد تكون X_k مرتبطة ببعض المتغيرات المستقلة أو بها كافة وإن كان هذا الارتباط ليس تاما. فإذا كان ذلك صحيحا فإنه يمكننا أن نفسر، جزئيا في الأقل، قيمة X_k بدلالة قيم المتغيرات المستقلة الأخرى. افترض أننا حاولنا عمل هذا بنموذج انحدار خطي يربط X_k بالمتغيرات المستقلة الأخرى:

$$X_{kt} = c_0 + c_1 X_{1t} + \cdots + c_{k-1} X_{(k-1)t} + v_{kt}, \qquad (4.15)$$

حيث u_{kt} هو الخطأ العشوائي. افترض أننا نرغب في تقدير المعلمات c_0, ..., c_{k-1} في المعادلة (4.15) عن طريق منهج المتغيرات المساعدة* والحصول على المقدرات $\hat{c}_0$ ، $\hat{c}_1$ ، ...، $\hat{c}_{k-1}$ باستخدام هذه المقدرات، تكون قيمتنا المفسرة (أو المحسوبة لـ X_k والمناظرة لـ X_1 ، ...، X_{k-1}) هي:

$$\hat{X}_{kt} = \hat{c}_0 + \hat{c}_1 X_{1t} + \cdots + \hat{c}_{k-1} X_{(k-1)t}, \qquad (4.16)$$

وطالما أننا افترضنا أنه X_{kt} ليست مولفا خطيا للمتغيرات المستقلة الأخرى، فإننا نعلم أنه عموما، أن $X_{kt} = \hat{X}_{kt}$. ولذلك يمكننا أن نضع X_{kt} في الشكل:

$$X_{kt} = \hat{X}_{kt} + \hat{v}_{kt}, \qquad (4.17)$$

حيث إن $\hat{v}_{kt}$ هي ذلك الجزء من X_{kt} والذي لايمكن أن ينسب إلى المتغيرات المستقلة الأخرى (أو $\hat{v}_{kt} = X_{kt} - \hat{X}_{kt}$) ويطلق على هذا الحد $\hat{v}_{kt}$ الباقي "The residual" في الانحدار الذي يربط X_k بالمتغيرات المستقلة X_1, ..., X_{k-1}.

وكما قد يخمن الفرد، فإن $\hat{v}_{kt}$ تكون مهمة جدا في تقدير b_k، لأنها تمثل الجزء من X_{kt} الذي يكون مستقلا، إلى حد ما، عن المتغيرات المستقلة الأخرى. فعلى سبيل المثال، من المعادلة (4.17) والمعادلة(4.16) نجد أنه، إذا كانت $\hat{v}_{kt}=0$ ، فإننا سنعاني وجود الارتباط الخطي المتعدد التام، ولن نستطيع تقدير b_k. وفي

* نشتق معادلتنا الطبيعية عن طريق وضع:

$$\Sigma \hat{v}_{kt} = 0, \Sigma\left(\hat{v}_{kt} X_{1t}\right) = 0, \cdots, \Sigma\left(\hat{v}_{kt} X_{k-1}\right) = 0.$$

هذه الحالة (وكما في الملحق)، أنه يمكننا أن نعبر عن المقدر لـb_k على النحو:

$$\hat{b}_k = \frac{\sum(\hat{v}_{kt} Y_t)}{\sum \hat{v}_{kt}^{\,2}} = \frac{\sum(X_{kt} - \hat{X}_{kt}) Y_t}{\sum (X_{kt} - \hat{X}_{kt})^2} \tag{4.18}$$

أي أن حل معادلاتنا الطبيعية لـ $\hat{b}_k$ يمكن التعبير عنه في الشكل (4.18) لاحظ أن قيم الباقي $\hat{v}_k$ تعتمد فقط على قيم X_1 إلى X_k، وهكذا، يمكن مشاهدتها. وينطبق القول نفسه على مقدراتنا الأخرى. وبالتحديد، نثبت في الملحق، أن:

$$\hat{b}_i = \frac{\sum(\hat{v}_{it} Y_t)}{\sum \hat{v}_{it}^{\,2}} = \frac{\sum(X_{it} - \hat{X}_{it}) Y_t}{\sum (X_{it} - \hat{X}_{it})^2}, \ i = 1, \cdots, k, \tag{4.19}$$

حيث إن $\hat{v}_{it}$ هو الباقي من انحدار X_i على المتغيرات المستقلة الأخرى كافة *. باختصار، يعتمد مقدرنا $(\hat{b}_i)$ على ذلك الجزء من X_i (وهو $\hat{v}_i$) غير المرتبط خطيا مع المتغيرات المستقلة الأخرى. وبالمقابل يمكننا الحصول على المقدر $(\hat{b}_i)$ عن طريق استعمال ذلك الجزء من X_i، فقط. الذي لايكون مولفا خطيا تاما مع المتغيرات المستقلة الأخرى كافة.

والآن دع $\hat{Y}_{ti}$ هو ذلك الجزء من Y_t المرتبط خطيا تاما مع المتغيرات المستقلة كافة باستثناء X_{it}* حينئذ، سيكون $(Y_t - \hat{Y}_{ti})$ هو ذلك الجزء من Y_t غير المرتبط خطيا

---

* لاحظ التشابه بين صيغة $\hat{b}$ في حالة الانحدار البسيط وحالة $\hat{b}_i$، في (4.19) ففي نموذج الانحدار البسيط في الفصل الثاني، أوضحنا أن:

$$\hat{b} = \frac{\sum(X_t - \overline{X}) Y_t}{\sum (X_t - \overline{X})^2}.$$

وعلى سبيل تدريب القارئ، عليه أن يثبت أن صيغة $\hat{b}$ في حالة الانحدار البسيط هي حالة خاصة من (4.19)؛ أي أنه، في حالة الانحدار البسيط، فإن $\hat{X}_t = \overline{X}$. [مساعدة في الحل: في حالة الانحدار البسيط تكون المعادلة المناظرة لـ(4.15) هي $X_t = c + v_t$.

بهذه المتغيرات المستقلة. بعد ذلك، نلاحظ أنه طالما أن $\hat{Y}_{ti}$ في المعادلة (4.19) على النحو:

$$\sum \hat{v}_{it} = 0, \sum (\hat{v}_i X_{jt}) = 0 \qquad j \neq i, \tag{4.20}$$

فلدينا بالضرورة:

$$\sum (\hat{v}_{it} \hat{Y}_{ti}) = 0. \tag{4.21}$$

ومن المعادلة (4.21)، وطالما أن $Y_t = \hat{Y}_{ti} + \left(Y_t - \hat{Y}_{ti}\right)$ ، يمكننا التعبير عن المقدر $\hat{b}_i$ في المعادلة (4.19) على النحو:

$$\hat{b}_i = \frac{\sum \hat{v}_{it} (Y_t - \hat{Y}_{ti})}{\sum \hat{v}_{it}{}^2} = \frac{\sum (X_{it} - \hat{X}_{it})(Y_t - \hat{Y}_{ti})}{\sum (X_{it} - \hat{X}_{it})^2}. \tag{4.22}$$

ينبغي أن يكون واضحا الآن كيف يعزل منهجنا تأثير المتغيرات المستقلة المختلفة على المتغير التابع Y_t. فالمقدر $\hat{b}_i$ في المعادلة (4.22) يعتمد، فقط، على قيم X_{it} و Y_t بعد استبعاد التأثير الخطي لجميع المتغيرات المستقلة الأخرى على كلا المتغيرين. في الحقيقة، يمكننا أن نتخيل أن منهجنا الحالي في التقدير يحول حالة المتغيرات المتعددة إلى حالة المتغيرين عن طريق استبعاد (طرح) تأثير المتغيرات الأخرى.

تباينات المقدرات

ينبغي عليك النظر لكيفية اشتقاق المعادلة (4.19) في ملحق هذا الفصل خاصة وأن المعالجة الدقيقة هناك تمكننا أولا: من إثبات أن $\hat{b}_i$ مقدر غير

---

* يمكن أن نحصل على ذلك عن طريق انحدار Y_t على المتغيرات المستقلة كافة باستثناء X_{it}، ثم حساب قيمته المتوقعة. على سبيل المثال، سنعتبر نموذج الانحدار:

$$Y_t = \gamma_0 + \gamma_1 X_{1t} + \cdots + \gamma_{i-1} X_{(i-1)t} + \gamma_{i+1} X_{(i+1)t} + \cdots + \gamma_k X_{kt} + w_t,$$

حيث w_t هو الخطأ العشوائي. بعدئذ، نستخدم طريقة المتغير المساعد للحصول على:

$\hat{\gamma}_0, \hat{\gamma}_1, \cdots, \hat{\gamma}_{(i-1)}, \hat{\gamma}_{(i+1)}, \cdots, \hat{\gamma}_k$ وحينئذ نحصل على:

$$\hat{Y}_{ti} = \hat{\gamma}_0 + \hat{\gamma}_1 X_{1t} + \cdots + \hat{\gamma}_{t-1} X_{(t-1)t} + \hat{\gamma}_{t+1} X_{(t+1)t} + \cdots + \hat{\gamma}_k X_{kt}.$$

متحيز لـ b_i (أي أن $E(\hat{b}_i)=b_i$). ثانيا: ايجاد صيغ للتباينات الشرطية لـ $\hat{b}_i$. ويمكننا، بالتحديد، أن نثبت:

$$\operatorname{var}(\hat{b}_i)=\sigma_{\hat{b}_i}^{\,2}=\frac{\sigma_u^{\,2}}{\sum \hat{v}_{it}^{\,2}}, \qquad for\ \ i=1,2,\cdots,k, \tag{4.23}$$

حيث σ_u^2 هو تباين الخطأ العشوائي في معادلة الانحدار المتعدد الأصلية (4.3). *

فترات الثقة واختبار الفرضيات: بعض المقدمات

بالاستمرار في افتراض، أن الأخطاء العشوائية موزعة توزيعا طبيعيا كما في حالة الانحدار البسيط، فسيكون صحيحا أيضا، - بافتراض قيم معطاة للمتغيرات المستقلة - أن المقدرات $\hat{b}_i$ كافة ستكون بدورها موزعة توزيعا طبيعيا. وينتج ذلك من كون $\hat{b}_i$ توليفة خطية من الأخطاء العشوائية. وكما ناقشنا في الفصل الثالث فإن التوليفات الخطية من المتغيرات الموزعة توزيعا طبيعيا تكون نفسها موزعة توزيعا طبيعيا. وهذا يمكننا من تجميع نتائجنا رمزيا على النحو: **

$$\hat{b}_i \quad is \quad N(b_i,\sigma_{\hat{b}_i}^{\,2}), \qquad\qquad i=0,1,\cdots,k, \tag{4.24}$$

حيث:

---

* لاحظ، مرة أخرى، التشابه بين صيغ التباين لـ $\hat{b}$ في حالة الانحدار ذي المتغيريـن و $\hat{b}_i$ في حالة الانحدار المتعدد. والفرق، مرة أخرى، هو أن الحد $\Sigma(X_t-\overline{X})^2$ في مقام صيغة التباين لـ $\hat{b}$ في حالة انحدار المتغيرين قد تم إحلاله بـ $\Sigma\hat{v}_{it}^2=\Sigma(X_{it}-\hat{X}_{it})^2$.

** المتغير المستقل المناظر لـ b_0 هو X_{0t} $[X_{0t}=1,\ t=1,2,\cdots,n]$. لذلك، سيكون $\hat{v}_{0t}$ هو الباقي من انحدار X_{0t} على المتغيرات المستقلة الأخرى كافة $X_{1t},\cdots,X_{kt}$.

$$X_{0t}=e_1X_{1t}+e_2X_{2t}+\cdots+e_kX_{kt}+v_{0t}$$

لاحظ أن معادلة الانحدار هذه لاتشتمل على حد ثابت، لأن الحد الثابت هو المتغير التابع. وعلى العكس، فإن معادلات الانحدار التي تعرف الأخطاء العشوائية v_{it}^{s} مثل المعادلة (4.15) تحتوي على الحدود الثابتة، وذلك بسبب أن واحداً من المتغيرات المفسرة هو المتغير التابع. باختصار، فإن X_{0t} ينبغي أن ينظر إليه على أنه متغير مفسر آخر.

$$\sigma_{\hat{b}_i}^{\ 2} = \frac{\sigma_u^{\ 2}}{\sum \hat{v}_{it}^2}$$

وتبقى الصعوبة أمامنا في أنه لايمكن تحديد تباينات مقدراتنا لأن σ_u^2 غير معلومة. ولذا، نحتاج لبناء مقدر لـ σ_u^2 من عينتنا من المشاهدات. وطالما أن σ_u^2 هي القيمة المتوسطة لـ u_t^2:

$$\sigma_u^{\ 2} = E[u_t^{\ 2}].$$

يكون من المنطقي، حينئذ، إذا علمنا قيم b_i، أن نأخذ متوسط تباينات العينة كمقدر لـ σ_u^2:

$$\frac{\sum u_t^{\ 2}}{n} = \frac{\sum (Y_t - b_0 - b_1 X_{1t} - \cdots - b_k X_{kt})^2}{n}. \tag{4.25}$$

ولسوء الحظ، فإننا لانعلم قيم المعلمات، إلا أنه لدينا مقدرات لها. ومن ثم، يمكننا إحلال كل من b_i في المعادلة (4.25) بمقدره، ومن ثم، الحصول على مقدر لتباين الخطأ العشوائي.

$$\hat{\sigma}_u^{\ 2} = \frac{\sum_{t=1}^{n} (Y_t - \hat{b}_0 - \hat{b}_1 X_{1t} - \cdots - \hat{b}_k X_{kt})^2}{n-(k+1)} \tag{4.26}$$

لاحظ أن المقام في المعادلة (4.26) هو [n - (k+1)] ويعكس هذا خسارة عدد (k+1) من درجات الحرية في البسط نتيجة تقدير معلمات عددها (k+1). وكما في حال انحدار المتغيرين، يكون صحيحا، أيضا:

$$E(\hat{\sigma}_u^{\ 2}) = \sigma_u^{\ 2}.$$

أي أن مقدرنا في المعادلة (4.26) غير متحيز. ويمكننا باستخدام المعادلة (4.26)، تقدير تباين كل من $\hat{b}_i$ بوساطة:

$$\hat{\sigma}_{\hat{b}_i}^{\ 2} = \frac{\hat{\sigma}_u^{\ 2}}{\sum \hat{v}_i^{\ 2}}. \tag{4.27}$$

فترات الثقة واختبار الفرضيات

نحن الآن في وضع يمكننا من بناء فترات الثقة للمعاملات المفردة. وباستخدام نتائجنا من الفصل الثالث، نلاحظ أنه طالما أن $\hat{b}_i$ متغير طبيعي $\left[\hat{b}_i \sim N\left(b_i, \sigma^2_{\hat{b}}\right)\right]$ فسيكون:

$$\frac{\hat{b}_i - b_i}{\sigma_{\hat{b}_i}}$$

متغيرا طبيعيا معياريا $N(0.1)$، حينئذ وبالطريقة نفسها التي اتبعناها في حالة الانحدار البسيط يمكننا إنشاء فترات الثقة، واختبار الفرضيات بدلالة المنحنى الطبيعي إذا كان σ^2_u (ومن ثم $\sigma^2_{\hat{b}_i}$) معلوما. على سبيل المثال ينبغي أن يثبت القارئ لنفسه أن فترة الثقة 95% لـ b_i، مبنية على المنحنى الطبيعي، هي:

$$\hat{b}_i \pm 1.96\,\sigma_{\hat{b}_i}\,.$$

وكما في حالة انحدار المتغيرين، تكون u^2 (عموما) غير معلومة، ولذا ينبغي تقديرها. ونتيجة لذلك ننشئ فترات الثقة أو نوجد نسب t لاختبار الفرضيات بدلالة توزيع t. وللقيام بذلك نلاحظ أن:

$$\frac{\hat{b}_i - b_i}{\hat{\sigma}_{\hat{b}_i}} \tag{4.28}$$

هو متغير t بدرجات حرية (n-k-1). لاحظ أن درجات الحرية للمتغير t في المعادلة (4.28) تكون مساوية دائما مقام مقدر التباين في المعادلة (4.26). وعلى سبيل المقارنة تذكر أنه، في حالة الانحدار البسيط، تكون k=1 ولذا (مع وجود معلمتين ينبغي تقديرهما يكون لدينا توزيع t بدرجات حرية قدرها (n-2).

يمكننا الآن، باستخدام المعادلة (4.28)، أن نتبع المنهج نفسه الذي طورناه في الفصل الثالث لاختبار الفرضيات المرتبطة بأي من المعاملات المفردة. على سبيل المثال، افترض أنه يوجد لدينا النموذج التالي:

$$Y_t = b_0 + b_1 X_{1t} + \cdots + b_9 X_{9t} + u_t, \qquad t = 1, \cdots, 25, \tag{4.29}$$

ونرغب في اختبار الفرضيات (عند 0.05 حجم النوع الأول من الخطأ):

$$H_0 : b_3 = 0,$$
$$H_1 : b_3 \neq 0,$$

وتتوافر لدينا عينة مكونة من 25 مشاهدة.

في هذه المسألة يكون لدينا n=25 و k=9 ومن ثم:

$$\frac{\hat{b}_3 - b_3}{\hat{\sigma}_{\hat{b}_3}}$$

يكون متغير t بدرجات حرية عددها (25-9=15). ومن الجدول الإحصائي ٢ لتوزيع t نجد أن فترة الثقة 95% لـ b_3 هي:

$$(\hat{b}_3 \pm 2.131\hat{\sigma}_{b_3}). \qquad (4.30)$$

وبالاستمرار، كما فعلنا في الفصل الثالث، يمكننا استخدام عينتنا لتقويم $\hat{\sigma}_{\hat{b}}, \hat{b}_3$ وبالتعويض عن هذه القيم في المعادلة (4.36)، ومعرفة ما إذا كانت الفترة الناتجة تغطي القيمة المفترضة الصفرية أم لا. فإذا كانت تغطيها فسوف نقبل H_0، أما إذا لم تكن تغطيها فسوف نرفضه. وبالمقابل يمكننا، ببساطة، أن نقيم نسبة t وباتباع قاعدتنا التجريبية، نقارن نسبة t المطلقة مع ٢. وفي حالتنا هذه تكون القيمة الحرجة الدقيقة هي 2.131، وعلى أي حال ينبغي أن يكون واضحا أنه عند اللحظة التي نعطي فيها نتيجة مثل المعادلة (4.28) تُحلُّ المشاكل المرتبطة باختبار فترات الثقة وتكوينها في حالة الانحدار المتعدد بالطريقة نفسها التي استخدمناها في حال انحدار المتغيرين.

(٤-٤) معامل التحديد المتعدد

بينا في المباحث السابقة طرق تقدير الفرضيات المرتبطة بالمعاملات الفردية في نموذج الانحدار المتعدد واختبارها. تبقى قضية إضافية متعلقة بالقوة التفسيرية لمعادلة الانحدار ككل. ما المقدار الذي يمكن تفسيره من التغير في المتغير التابع عند أخذ المتغيرات المستقلة كافة معا؟

R^2 لحالة الانحدار المتعدد

أوجدنا في الفصل الثاني مقياس R^2 لحالة الانحدار البسيط. تذكر أن R^2 تحتوي على قيمة بين الصفر والواحد الصحيح تشير إلى جزء التغير في Y الذي يمكن أن تفسره معادلة الانحدار المقدرة. وباتباع المنهج نفسه فسوف نثبت أن القاعدة نفسها R^2 بالتفسير نفسه قابلة للتطبيق، أيضا، لحالة الانحدار المتعدد.

اعتبر نموذجنا الأساسي للانحدار المتعدد التالي:

$$Y_t = b_0 + b_1 X_{1t} + \cdots + b_k X_{kt} + u_t . \tag{4.31}$$

لاحظنا سابقا أنه إذا قدرت المعادلة (4.31) بوساطة طريقة المتغير المساعد فإنه يمكن التعبير عن Y_t على النحو:

$$Y_t = \hat{Y}_t + \hat{u}_t , \tag{4.32}$$

حيث:

$$\hat{Y}_t = \hat{b}_0 + \hat{b}_1 b_{1t} + \cdots + \hat{b}_k X_{kt} , \tag{4.33}$$

و $\hat{u}_t$ حيث تكون:

$$\sum \hat{u}_t = 0, \quad \sum (\hat{u}_t X_{it}) = 0, \qquad i = 1, \cdots, k \tag{4.34}$$

والآن نجمع المعادلة (4.32) على مدى عينتنا لنحصل على:

$$\sum Y_t = \sum \hat{Y}_t + \sum \hat{u}_t . \tag{4.35}$$

ولما كانت $\Sigma \hat{u}_t = 0$ ، يكون لدينا كما في حالة الانحدار البسيط

$$\sum Y_t = \sum \hat{Y}_t . \tag{4.36}$$

وبقسمة حدود المعادلة (4.36) على n يكون متوسط العينة لـ Y_t يساوي متوسط العينة لـ $\hat{Y}_t$:

$$\bar{Y} = \bar{\hat{Y}} . \tag{4.37}$$

وبالعودة إلى المعادلة (4.32) وتربيع جانبي المعادلة:

$$Y_t^2 = \hat{Y}_t^2 + 2 \hat{Y}_t \hat{u}_t . \tag{4.38}$$

وبالجمع على مدى العينة، نحصل على:

$$\sum Y_t^2 = \sum \hat{Y}_t^2 + \sum \hat{u}_t^2 + 2\sum (\hat{Y}_t \hat{u}_t). \quad (4.39)$$

ويساوي الحد الأخير في المعادلة (4.39) الصفر، ولنرى ذلك لاحظ أن:

$$\sum (\hat{u}_t \hat{Y}_t) = \sum \hat{u}_t (\hat{b}_0 + \hat{b}_1 X_{1t} + \cdots + \hat{b}_k X_{kt})$$

$$= \hat{b}_0 \sum \hat{u}_t + \hat{b}_1 \sum (\hat{u}_t X_{1t}) + \cdots + \hat{b}_k \sum (\hat{u}_t X_{kt}) = 0,$$

ويمكن في ضوء المعادلة (4.34) تبسيط المعادلة (4.39) إلى:

$$\sum Y_t^2 = \sum \hat{Y}_t^2 + \sum \hat{u}_t^2. \quad (4.40)$$

نطرح بعد ذلك $n\overline{Y}^2$ من جانبي المعادلة (4.40):

$$\sum Y_t^2 - n\overline{Y}^2 = (\sum \hat{Y}_t^2 - n\overline{Y}^2) + \sum \hat{u}_t^2. \quad (4.41)$$

وبتذكر أنه من (4.31) أن $\overline{\hat{Y}} = \overline{Y}$، يمكننا أن نكتب المعادلة (4.41) على النحو:*

$$\sum (Y_t - \overline{Y})^2 = \sum (\hat{Y}_t - \overline{Y})^2 + \sum \hat{u}_t^2, \quad (4.42)$$

والتي تكون متماثلة مع المعادلة المناظرة لحالة الانحدار البسيط والتي تم اشتقاقها في الفصل الثاني.

تذكر أننا وضعنا هذه العلاقة على النحو:

$$\text{TSS} = \text{RSS} + \text{ESS} \quad (4.43)$$

حيث إن $\text{TSS} = \Sigma(Y_t - \overline{Y})^2$ و $\text{RSS} = \Sigma(\hat{Y}_t - \overline{Y})^2$ وأخيرا $\text{ESS} = \Sigma\hat{u}_t^2$. ويعبر المجموع الكلي للمربعات TSS عن التغير في المتغير التابع حول الوسط الحسابي للعينة الذي نحاول تفسيره بمعادلتنا للانحدار. أي أنه يفترض أن يخبرنا نموذجنا للانحدار لماذا تكون Y_t غير ثابتة. وذلك الجزء الذي لايمكن للمعادلة أن تفسره هو مجموع مربعات الخطأ (ESS). والفرق بين TSS و ESS ينبغي أن يكون ذلك الجزء الذي تفسره معادلة الانحدار وهو (RSS) مجموع مربعات الانحدار.

* نستخدم هنا إحدى قواعد الفصل الأول التي تعني أنه بالنسبة لأي متغير Z_t:

$$\sum (Z_t - \overline{Z})^2 = \sum Z_t^2 - n\overline{Z}^2.$$

وكما هو في حالة نموذج الانحدار ذي المتغيرين، نستخدم الجزء من المتغير في القيم المشاهدة لـ Y الذي يمكن أن ينسب إلى معادلة الانحدار المقدرة بوصفه مقياسا للقوة التفسيرية لمعادلة الانحدار.

$$R^2 = 1 - \frac{ESS}{TSS} = \frac{RSS}{TSS} = \frac{\sum(\hat{Y}_t - \bar{Y})^2}{\sum(Y_t - \bar{Y})^2}, \tag{4.44}$$

حيث يطلق على R^2 معامل التحديد المتعدد.

باختصار، إذا كان لدينا توفيق تام فإن القيمة المحسوبة تتساوى في كل حالة مع القيم المشاهدة $\left(Y_t = \hat{Y}_t\right)$ أي أن $\hat{u}_t$ تساوي الصفر وأن ESS=0، لذلك يكون RSS=TSS، ويحقق R^2 قيمته العظمى بأن تساوي الواحد الصحيح. وبالمقابل، إذا كانت المعادلة المقدرة لاتفسر أيا من التغير في المتغير التابع فسيكون ESS أكبر مايمكن ويتعادل هنا مع TSS (ESS=TSS). ولذا، يكون RSS=0. وهنا يكون $R^2 = 0$ وكلما اقترب R^2 من الواحد ازدادت القوة التفسيرية لمعادلة الانحدار المقدرة. وأخيرا، وعلى الرغم من أننا لن نعطي إثباتا (طالما أن الحالة متماثلة مع حالة انحدار المتغيرين) نشير إلى أن R قد لاتعني أكثر من مجرد مقدر لمعامل الارتباط بين Y_t و $\hat{Y}_t$.

تعليق على R^2

لما كان $R^2 = 1 - ESS/TSS$ معامل التحديد، فإنه يعتمد بوضوح على كل من ESS و TSS. ومن بين هذين المحددين لـ R^2، ترتبط ESS بالقوة التفسيرية للنموذج. ذلك أن TSS، ببساطة، هو مقياس لتغير العينة في المتغير الذي نحاول تفسيره، وهو بذلك لايعتمد على سمات النموذج المستخدم لتفسير Y_t. وهكذا يرتبط R^2 بالقوة التفسيرية للنموذج بحكم ارتباط ESS بها.

اعتبر الآن الحالة التي يرغب فيها اثنان من الباحثين في تفسير قيم عددها n من Y_t، $t=1,2,...,n$. افترض أن الباحث الأول يستخدم النموذج (4.31)، أما الباحث الثاني فيستخدم النموذج التالي:

$$Y_t = b_0 + b_1 X_{1t} + \cdots + b_k X_{kt} + b_{k+1} X_{k+1,t} + \cdots + b_{k+r} X_{k+r,t} + u_t . \qquad (4.45)$$

أي أن الباحث الثاني يدخل في نموذجه المتغيرات المستقلة كافة التي اعتبرها الباحث الأول، إضافة إلى بعض المتغيرات الأخرى (أي X_{k+1t} ، . . . ، X_{k+rt}). دع ESS_1 وRSS_2 هما مجموع مربعات الخطأ التي حصل عليها بوساطة الباحث الأول والباحث الثاني على الترتيب. حينئذ، وكما سنبين أدناه، فإن:

$$ESS_2 \le ESS_1, \qquad (4.46)$$

ونتساءل الآن هل المتغيرات الإضافية التي اعتبرها الباحث الثاني ملائمة ! دع R_1^2 وR_2^2 وهما معاملي التحديد اللذين حصل عليهما الباحثان الأول والثاني على الترتيب. وبما أن$R_1^2 = 1 - ESS_1 / TSS$ و$R_2^2 = 1 - ESS_2 / TSS$ فإنه ينتج من المعادلة (4.46)أن تكون:

$$R_1^2 \le R_2^2 . \qquad (4.47)$$

مرة أخرى، تظل المعادلة (4.47) صحيحة سواء كانت المتغيرات المستقلة الإضافية ملائمة أم لا. ويظل عدم التساوي في المعادلة (4.46) والمعادلة (4.47) صحيحا دائما. ولكننا نلاحظ في التطبيقات العملية، عادة، أنه إذا لم يكن$ESS_1 = 0$ و $R_1^2 = 1$ فإن:

$$ESS_2 \le ESS_1, \text{ و } R_2^2 \rangle R_1^2 \qquad (4.48)$$

وتشير هذه النتيجة إلى أن إحصائية R^2 تستمر في التزايد ولن تنخفض أبدا بإضافة متغيرات مستقلة إضافية إلى النموذج سواء كانت هذه المتغيرات ملائمة أم لا. والسبب في ذلك هو، وكما أوضحنا في ملحق الفصل الثاني، أن مقدراتنا في حالة الانحدار ذي المتغيرين والناتجة عن استخدام طريقة المتغيرات المساعدة متطابقة مع المقدرات الناتجة باستخدام طريقة المربعات الصغرى. ويكون هذا التطابق صحيحا، أيضا، في حالة الانحدار المتعدد. أي أن مقدراتنا للمعلمات الناتجة عن استخدام طريقة المتغيرات المساعدة لنموذج الانحدار المتعدد الخطي متطابقة مع المقدرات التي يحصل عليها باستخدام طريقة المربعات الصغرى.

دع $\hat{b}_0$، ...، $\hat{b}_k$ هي تقديرات معاملات الانحدار التي حصل عليها الباحث الأول. حينئذ تكون هذه التقديرات بشكل يكون معه مجموع مربعات الخطأ:

$$ESS_1 = \sum_{t=1}^{N}\left(Y_t - \hat{b}_0 - \hat{b}_1 X_{1t} - \cdots - \hat{b}_k X_{kt}\right)^2 \tag{4.49}$$

صغيرا بقدر الإمكان. وينتج هذا من التطابق بين طريقتي المتغير المساعد والمربعات الصغرى.

دع الآن $\tilde{b}_0$، ...، $\tilde{b}_k$، $\tilde{b}_{k+1}$ و $\tilde{b}_{k+r}$ هي التقديرات التي حصل عليها الباحث الثاني. حينئذ ومرة أخرى وبسبب التماثل مع طريقة المربعات الصغرى، تكون هذه التقديرات بشكل يصبح معه ESS_2 صغيرا بقدر الإمكان عندما يكون:

$$ESS_2 = \sum_{t=1}^{N}(Y_t - \tilde{b}_0 - \cdots - \tilde{b}_k X_{kt} - \tilde{b}_{k+1} X_{k+1,t} - \cdots - \tilde{b}_{k+r} X_{k+r,t})^2. \tag{4.50}$$

وفي هذه الحال، سيتساوى ESS_2 مع ESS_1. هو أن إحدى المجموعات الممكنة من القيم لـ هي:

$$\tilde{b}_i = \hat{b}_i, \qquad i = 0, \cdots, k \tag{4.51}$$

$$\tilde{b}_{k+j} = 0, \qquad j = 1, \cdots, r.$$

وبسبب أن ESS_2 لن يزيدا ابدا على ESS_1، ينتج عن ذلك أن القوة التفسيرية للنموذج الثاني (كما تقاس بوساطة R^2) لن تكون اقل من نظيرتها في النموذج الأول ابدا. ولكن يحصل على قيم أخرى لـ $\tilde{b}_1$، ...، $\tilde{b}_k$ ولذا يكون $ESS_2 < ESS_1$ ومن ثم، يكون $R_2^2 > R_1^2$.

معامل التحديد المعدل $\overline{R}^2$

يرغب الباحثون، عادة، في مقارنة نماذجهم بدلالة مختلف المقاييس لمدى جودة "goodness" النموذج . وتوحي مناقشتنا أن إحصائية R^2 يمكن أن تكون واحدة من هذه المقاييس ولكن مع الحذر. ذلك أنه إذا كانت R^2 تزداد، عادة، بإضافة

متغيرات مستقلة جديدة في النموذج (بغض النظر عن مدى ملائمتها لـه)، فـإن الباحث يمكنه الوصول إلى نموذج ممتاز يتفوق على النماذج الأخرى باستخدام هذا المقياس عن طريق إضافة متغيرات مستقلة جديدة.

وتكمن الصعوبة التي نواجهها مع إحصائية R^2 في عدم وجود جزاء متر تب على زيادة عدد المتغيرات المستقلة المستخدمة في تفسير المتغير التابع. ولذا، نستخدم في بعض الاحيان شكلا معدلا مـن R^2. ويتضمن هذا الشكل الجزاء المشار إليه أعلاه.

يرمز، عادة، لمعامل التحديد المعدل بالرمز $\overline{R}^2$. عموما افترض أننا نأخذ نموذج الانحدار الخطي بحد ثابت، وبعدد p من معاملات الانحدار كعدد اجمالي، فعلى سبيل المثال، بالنسبة لـ المعادلة (4.31) تكون p=k+1 وبالنسبة للمعادلة (4.45) تكون p=k+r+1 حينئذ، يكون معامل التحديد المعدل هو:

$$\overline{R}^2 = 1 - \frac{ESS/(n-p)}{TSS/(n-1)} \tag{4.52}$$

حيث إن n هي حجم العينـة، ESS و TSS هما مجموع مربعات الخطأ ومجموع المربعات الكلية للنموذج موضع الاعتبار على الترتيب.*

ويمكن آن تزداد قيمـة $\overline{R}^2$، أو أن تظل على حالها، أو تنقص إذا أضفـنـا متغيرات مستقلة اضافية للنموذج، وسبب ذلك هو أن مثل هذه المتغيرات المستقلة الإضافية سوف تخفض عادة من قيمة ESS، ولكنها سوف تقلل، أيضا، من n-p. ولذا لايمكن التنبؤ باتجاه التغير في ESS/(n-p)، ومن ثم، اتجاه التغير في $\overline{R}^2$.

ويشعر الباحثون، غالبا، أنه إذا أضيف متغير مهم أو مجموعة متغـيـرات مهمة للنموذج فإن النقص في ESS سيفوق الحجم اللازم للتعويض عن النقص في (n-p) ولذا ستزداد $\overline{R}^2$. بالطبع، ستكون هذه النتيجة صحيحة لبعض المتغيـرات

* باستخدام المعادلة (4.52) والمعادلة (4.44)، يمكن أثبات أن $\overline{R}^2 \le R$ إذا كانت $R^2 = 1$، حينئذ، فإن $R^2 = \overline{R}^2$، ولكن، إذا كانت $R^2 < 2$ و $p \ge 2$ فإن $\overline{R}^2 < R^2$ للمهتمين بهذا الموضوع، نقدم إثباتاً لهذه النتائج في الملحق B في ختام هذا الفصل

المهمة ولكننا نؤكد هنا أن هذه ليست نتيجة نظرية دقيقة لأنه يمكن إيجاد أمثلة مناقضة لها. فعلى سبيل المثال في المعادلة (4.45)، افترض أن $r=1$ وأن $X_{k+1,t}$ متغير مهم لأن معامله لايساوي الصفر أي $b_{k+1} \neq 0$. في هذه الحال، يعتمد المتغير التابع Y_t، في الحقيقة، على $X_{k+1,t}$ (بالإضافة إلى المتغيرات الأخرى). على الرغم من هذه الأهمية، يمكن أن يظل $\overline{R}_1^2 < \overline{R}_2^2$ حيث إن $\overline{R}_1^2$ هو معامل التحديد المعدل المناظر للمعادلة (4.45) وأن $\overline{R}_1^2$ المناظر للمعادلة (4.31).

ولما كان كثير من الباحثين يشعر أن $\overline{R}^2$ سيتزايد إذا أضيفت متغيرات «مهمة» وينقص إذا أضيفت متغيرات «غير مهمة» فإن إحصائية $\overline{R}^2$ تستخدم للمقارنة بين النماذج افترض، على سبيل المثال، أن النماذج موضع الاعتبار هي المعادلة (4.45) والمعادلة (4.31). حينئذ، سيأخذ بعض الباحثين النموذج «الحقيقي» أو «الأفضل» على أنه ذلك النموذج الذي يحتوي على $\overline{R}^2$ أكبر.

وعلى الرغم من الاغراءات الحدسية لهذا النموذج، إلا أنه لايستند على أسس علمية، ولذا، فإننا لانوصي باستخدامه. وسبب ذلك هو أن مثل هذه المقارنات هي شكل من اشكال اختيار صدق أحد النماذج أوصحته إزاء النموذج الآخر، إلا أن خصائص هذا الاختبار، كأخطائه من النوع الأول ومن النوع الثاني، غير معلومة. لذلك، وفي الواقع، فإن الباحثين الذين يقيمون النماذج بدلالة احصائية $\overline{R}^2$ بهذه الطريقة المباشرة أنما يستخدمون اختبارا لم تعرف خصائصه بعد.

أما منهج الاختبار الملائم لتقويم المعادلة (4.45) بالنسبة للمعادلة (4.31) فيكون على النحو التالي: أولا، اعتبر $r=1$. في هذه الحال، يكون اختبار فرضية العدم الملائم $H_0 : b_{k+1} = 0$. ومن الواضح أنه إذا كانت $b_{k+1} = 0$ فإن المعادلة (4.31) تكون محددة تحديدا صحيحا، وإذا كانت $b_{k+1} \neq 0$ غير محددة تحديدا صحيحا بينما تكون المعادلة (4.45) محددة بصورة صحيحة. افترض لتبسيط الطرح، أن الفرضية البديلة هي $H_1 : b_{k+1} \neq 0$. افترض، أخيرا، أننا نرغب في جعل الخطأ من النوع الأول مساويا لـ 0.05 حينئذ، مع تحقق افتراضات النموذج الموجودة في المبحث (٤-١)،

سوف يتم رفض H_0 لصالح H_1 إذا كانت $\left|\hat{b}_{k+1} / \hat{\sigma}_{\hat{b}_{k+1}}\right| > t_{n-k-2\ 0.975}$ حيث إن $\hat{b}_{k+1}$ هو مقدرنا لـ b_{k+1} وفقا لطريقة المتغير المساعد، و $\hat{\sigma}_{\hat{b}_{k+1}}$ هو المقدر المناظر للانحراف المعياري لـ $\hat{b}_{k+1}$، وقد ناقشتنا طريقة هذا الاختبار في المبحث (٤-٣).

اعتبر الآن $r > 1$، في هذه الحال، تكون فرضية العدم موضع الاهتمام هي:

$$H_0 : b_{k+1} = b_{k+2} = \cdots = b_{k+r} = 0. \tag{4.53}$$

من الواضح أنه إذا كانت H_0 في المعادلة (4.53) صحيحة، فإن النموذج (4.31) يكون محددا تحديدا صحيحا. ولقد شرحت إحدى الطرق لاختبار الفرضيات من النوع الموجود في المعادلة (4.53) مقابل الفرضية البديلة ذات الطرفين في الملحق B من الفصل الخامس. مرة أخرى، وعلى العكس من اختبار $\overline{R}^2$ يمكن من خلال الطريقة -الموجودة في الملحق B من الفصل الخامس، وضع حجم الخطأ من النوع الأول عند المستوى المرغوب فيه (مثلا 0.05). خلاصة القول هو أنه لاينبغي تفضيل نموذج على آخر بسبب أنه يتضمن قيمة أعلى لـ $\overline{R}^2$.

وقد يحتار القارئ في السبب الذي فعلناه لتضمين المناقشة الخاصة بمعامل التحديد المعدل أو $\overline{R}^2$ في الوقت الذي نوصي بقوة بعدم استخدامه. ان المنطق وراء ذلك مزدوج. أولا، كثيرا ماتطبع برامج الحاسوب قيم R^2 و $\overline{R}^2$. فإذا اهملنا $\overline{R}^2$ فقد تتعجب مإذا تعني هذه الإحصائية. وثانيا، وكما أشرنا، فإن بعض الباحثين يقارنون مدى جودة التوفيق لنماذجهم من خلال مقارنة قيم $\overline{R}^2$. ونحث مرة، أخرى، على الحذر من قبول مثل هذا الاختبار، لأنه، كما أشرنا من قبل، لايستند إلى أسس علمية صحيحة.

وهناك نقطة أخيرة هي أننا نلاحظ أن هناك منهجا علميا لاختبار الفرضيات من النوع الموجود في المعادلة (4.53) المبني على مقارنات للإحصائية $\overline{R}^2$، وفي الواقع، فإن هناك اختبار آخر مبنيا على مقارنة إحصائيات R^2. ولكن هذين المنهجين، في الحقيقة، يمثلان المنهج الشائع الاستخدام والموصوف في الملحق B من الفصل الخامس. ولما كانت الاعباء الحسابية لهذه الاختبارات ليست أقل من الاختبارات

الموصوفة في الملحق B، وطالما أنها لاتقدم لنا إضافات جديدة فإننا لن نشرحها في هذا الكتاب.

(٤-٥) تحليل الانحدار المتعدد: مثالان توضيحيان

بعد أن درسنا مبادئ تحليل الانحدار، سنفحص الآن دراستين فعليتين للانحدار المتعدد لنتعرف على المنهج المستخدم ولنفسر النتائج الخاصة بهاتين الدراستين.

دالة استهلاك متعددة المتغيرات

افترضنا، في بداية هذا الفصل، أن حجم الإنفاق الاستهلاكي قد يعتمد على عدد من المتغيرات إضافة إلى مستوى الدخل الحالي. وفي الحقيقة، كون الاقتصاديون عددا من نظريات الاستهلاك وأجروا اختبارات قياسية مكثفة لهذه النظريات. فعلى سبيل المثال، طرح البرت اندو - وفرانكو مودجلياني نظرية دورة الحياة للاستهلاك حيث يريان أن مستوى الاستهلاك الحالي لأحد الأفراد يعتمد على القيمة المتوقعة لتيار دخله المستقبلي مدى الحياة.* وعلى سبيل الاختبار العملي البسيط لهذا الافتراض، اقترحا أن:

$$C_t = b_0 + b_1 Y_{dt} + b_2 A_{t-1} + u_t, \tag{4.54}$$

حيث يستعمل دخل العمل المتاح الحالي(Y_{dt}) متغيرا تقريبيا عن الدخل المتوقع من خدمات العمل، وA_{t-1} هو القيمة الصافية لثروة المستهلكين في نهاية الفترة (t-1) مقياس لدخل الملكية المتوقع. أي أن المستهلكين يبدأون الفترة t بثروة صافية قدرها A_{t-1} والتي تزودهم بالدخل الريعي أو الفائدة. قدر أندو ومودجلياني بعد ذلك قيم المعاملات في المعادلة (4.54) باستخدام تحليل الانحدار المتعدد وباستخدام بيانات سنوية عن الإنفاق الاستهلاكي، دخل العمل المتاح، وصافي الثروة، وقيست جميعا ببلايين الدولارات الجارية للسنوات ١٩٢٩-١٩٥٩م (بعد استبعاد سنوات

* انظر: "The Life Cycle Hypothesis of Saving : Aggregate Implications and Tests." *American Economic Review*, 53(March 1963), pp. 55-84.

الحرب ١٩٤١-١٩٤٥م). وجمعا ٢٥ مشاهدة مشتركة لهذه المتغيرات، ثم استخدما بيانات هذه السلسلة الزمنية في نموذج الانحدار المتعدد لتقدير المعادلة (4.54) ووجدا أن:

$$\hat{C}_t = 5.33 + 0.767 Y_{dt} + 0.047 A_{t-1} \qquad N = 25 \qquad (4.55)$$

$$(1.46) \quad (0.047) \quad (0.010) \qquad R^2 = 0.999$$

(ملاحظة: الأرقام الموجودة داخل الأقواس أسفل المعاملات المقدرة هي قيم الأخطاء المعيارية المناظرة).

من أول الأشياء التي نلاحظها في نتائج اندو - مودجلياني هي أن المعاملات المقدرة، وكما توقعنا موجبة، وأن الميل الحدي للاستهلاك من دخل العمل المتاح يقع بين الصفر والواحد الصحيح، لذلك تبدو المعادلة المقدرة متسقة مع السمات المفترضة للسلوك الاستهلاكي.

دعنا نلاحظ، بعد ذلك، من المعادلة (4.55) أنه - في الحالات الثلاث جميعا - تزيد تقديرات المعلمة على ثلاثة أضعاف تقديرات الأخطاء العشوائية. وباستخدام قاعدتنا التجريبية المرتبطة بنسب t، يتضمن هذا أنه إذا اعتبرت أية فرضية من فرضيات العدم $b_2 = 0$ و $b_1 = 0$ و $b_0 = 0$ مقابل الفرضيات البديلة ذات الذيل الواحد أو ذات الذيلين فسوف ترفض عند مستوى معنوية 5% أو 1% على أساس نتائج المعادلة (4.55).

وعلى سبيل توضيح إضافي، افترض أننا نرغب في اختبار الفرضية بأن b_1، أو MPC من دخل العمل المتاح، تساوي 0.5 مقابل الفرضية البديلة ذات الذيلين:

$$H_0 : b_1 = 0.5,$$
$$H_1 : b_1 \neq 0.5.$$

وبـ 22 درجة حرية ومستوى معنوية 5% نجد من الجدول الإحصائي رقم٢ أن القيمة الحرجة لـ t هي 2.07، ومن المعادلة (4.55) تكون القيمة المحسوبة لنسبة t هي:

$$\frac{0.767 - 0.5}{0.047} = \frac{0.267}{0.047} = 5.7 > 2.07.$$

وعليه، سوف نرفض فرضية العدم عند مستوى معنـويـة 5% على أساس نتائـج المعادلة (4.55)*. وفي الحقيقة، فإنه، عندما تكتسب بعض الخبرة في تحليل الانحدار، سوف تكتشف أن بإمكانك أداء كثير من هذه الاختبارات بوساطة الفحص الأولي. وعلى سبيل المثال، ففي الحالة السابقة، بلمحـة واحـدة لـ b_1 (= 0.767)، سوف يتضح لنا أنها تبعد عن القيمة المفترضـة لـ b_1 (أو 0.5) بأكثر من ضعفي الخطـأ العشوائي (2 x 0.046). ونتيجة لذلك، فإن الفرضية بأن $b_1 = 0.5$ يمكن رفضها بدون إجراء أي حسابات.

ونلاحظ، أخيرا، أن المعادلة التي كونها أندو وموديلياني لها قوة تفسيرية كبيرة. حيث إن $R^2 = 0.999$، وهذا يعني أنه يمكن أن ينسب للمتغيرين المستقلـين 99% من التغير في القيم المشاهدة للاستهلاك. فإذا قبلنا الشكل العملـي لـهـذه النظرية فإن هذه النتائج العملية تبدو متسقة مع نظريتهما للسلوك الاستهلاكي.

ومن المهم أن ندرك، عند تفسير هذه النتائج، امكانية أن تكون نتائج هذه المعادلة متسقة مع نظريات أخرى أيضا. ذلك أنه، بينما يمكننا التصريح بأن النتائج العملية تتفق مع النظرية، فإنه لايمكننا أن ندعي بأن نظريات الاستهلاك الأخرى غير صحيحة. ويبين هذا سبب كون التطبيق العملي على مشكلة معـيـنـة فـي الاقتصاد، عادة، عملية مستمرة، حيث تتراكم الأدلة التي تتفق مع سبب كـون افتراض معين أو تتعارض معه. ذلك أن الدعم التطبيقي لفرضية ما يعتمد علـى المدى الذي تتسق فيه النتائج مع الفرضية. وبالدرجة نفسها من الأهمية على المدى الذي تكون فيه هذه النتائج غير متسقة مع الفرضيات المنافسة.

دراسة لضرائب المدينة

لختم هذا الفصل، ستفحص دراسة أخرى عن الانحدر المتعدد، في هـذه المرة باستخدام البيانات المقطعية. تظهر في هذه الأيام بعض المشاكل الحـضـريـة

* ينبغي أن يكون واضحا أن H_0 سوف ترفض، أيضا، عند مستوى 1% من المعنوية.

المهمة، ويبحث رؤساء المدن عن مصادر للإيرادات الاضافية لتغطية نفقاتهم المتزايدة. واصبحت المشكلة تتلخص (إلى حد كبير) في محاولة فرض ضرائب جديدة أو رفع معدلات الضرائب القائمة دون إحداث خسارة مهمة في اقتصاد المدينة، خاصة أن هناك اقتناعا لدى كثير من المراقبين أن المعدلات العالية من الضرائب قد عجلت من نزوح الطبقة المتوسطة إلى الضواحي، كما قللت شراء المستهلكين حاجياتهم من السلع والخدمات من المدن. أحد مصادر الإيرادات الرئيسة للمدن في الولايات المتحدة الأمريكية هي الضريبة على مبيعات التجزئة. يقال إن مثل هذه الضرائب تؤدي إلى انخفاض في مبيعات التجزئة وقد ساهمت في تدهور اقتصاديات المدن.

هل هذا صحيح، وإذا كان كذلك، هل له أهمية كمية كبيرة؟ وهذا، بالطبع، سؤال تصعب الاجابة عنه، ولكن دعنا نفكر بعض الشئ في كيفية استخدام التحليل القياسي لتناول هذه المشكلة، إذا أدت ضرائب المبيعات الأعلى في وسط المدن مقارنة بالضرائب في الضواحي إلى انخفاض المبيعات في المدينة. فسوف نتوقع أن نجد - مع بقاء الأشياء الأخرى على حالها - أنه، كلما كانت الاختلافات أكبر بين معدل الضرائب على المبيعات في مدينة معينة وبين الضرائب في ضواحيها، فسيقل معدل المبيعات لكل نسمة في المدينة عن نظيرها في الضواحي. فإذا استطعنا الحصول على عينة من المدن تتوافر عنها بيانات عن مبيعات التجزئة لكل نسمة وحجم هذه الاختلافات الضريبية، فإنه قد يمكننا أن نكون انحدارا لمتغير المبيعات على متغير الاختلافات في المعدلات الضريبية، ثم نقوم، بفحص، بعد ذلك، العلاقة وحجم المعامل المقدر والقيام بالاختبارات اللازمة.

ويبدو هذا المنهج معقولا بدرجة كافية، إلا أن فيه مشكلة متأصلة تجعلنا لانشعر بالارتياح. ذلك أن فرضيتنا هي، مع بقاء الأشياء الأخرى على حالها، ارتباط اختلافات ضرائب المبيعات في المدن مع مستوى أقل من المبيعات. ولكن، في حدود عينتنا من المدن فإن الاشياء الأخرى لاتظل على حالها. ذلك أن مبيعات التجزئة في مدينة معينة تعتمد على عدد من المتغيرات المهمة إضافة إلى الضرائب،

كمستوى الدخل والحجم النسبي لسكان الضواحي. على سبيل المثال، إذا كان قاطنوا المدينة من الأغنياء يتوقع أن تكون المبيعات لكل نسمة أعلى، إضافة إلى ذلك، كلما ازداد عدد المشترين المحتملين في المدينة فستكون مبيعاتها أكبر. ولذا، فإن ماينبغي علينا عمله هو التحكم في تأثير هذه المحددات الأخرى لمبيعات المدينة حتى يمكننا عزل تأثير الاختلافات في ضريبة المبيعات أو فصلها.

ويستدعي هذا، بالطبع أن نستخدم تحليل الانحدار المتعدد، وهنا يمكننا، بالتحديد، أن نكون انحدارا لمتغير المبيعات على مجموعة من المتغيرات المستقلة التي ستشتمل ليس فقط على متغير الضريبة ولكن أيضا على محددات مهمة للمبيعات في المدينة. مثل هذه الدراسة نفذهاجون ميكسل John Mikesell*، والذي اقترح النموذج:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + b_3 X_{3t} + b_4 X_{4t} + u_t, \tag{4.56}$$

حيث:

Y_t = نصيب الفرد من مبيعات التجزئة في المدينة t.

X_{1t} = الاختلافات في ضريبة المبيعات للمدينة t.

X_{2t} = الدخل لكل نسمة في المدينة t.

X_{3t} = نسبة سكان المدينة t إلى اجمالي السكان في المدينة وضواحيها.

X_{4t}= مساحة المدينة t (بالأميال المربعة).

وقد عرف متغير الاختلافات في ضريبة المبيعات على النحو:

$$X_{1t} = \frac{1+t_c}{1+t_s},$$

حيث إن t_c هو معدل الضريبة على المبيعات في المدينة و t_s هو معدل الضريبة المتوسط على المبيعات في البلديات المحيطة بها. لذلك، فزيادة معدل ضريبة

---

* "Central Cities and Sales Tax Rate Differentials : The Border City Problem." *National Tax Journal* 23 (June 1970), pp. 206-213

المبيعات في المدينة t مع عدم تغير t_s سوف يزيد X_{1t} وسيخفض طبقا للمعادلة (4.56) المبيعات لكل نسمة، طالما أننا نفترض $b_1 < 0$.

استطاع ميكسل أن يجمع بيانات لعينة من 173 مدينة رئيسة وضواحيها في الولايات المتحدة الأمريكية، واستخدم هذه البيانات لتقدير المعادلة (4.56) باستخدام الانحدار المتعدد ووجد أن:

$$\hat{Y} = 4.5 - \underset{(2.94)}{7.44X_1} + \underset{(0.10)}{0.43X_2} - \underset{(0.04)}{0.11X_3} - \underset{0.02}{0.08X_4} \qquad \underset{R^2=0.26,}{N=173} \qquad (4.57)$$

حيث الأرقام داخل الأقواس هي الاخطاء المعيارية المقدرة المناظرة. ونلاحظ مباشرة أن معامل متغير الضريبة X_1 له العلاقة السالبة المتوقعة وأكبر من ضعف حجم الخطأ المعياري المناظر، ويعني ذلك أن نسبة t تزيد على ٢. يوحي هذا بوجود علاقة سالبة بين مبيعات التجزئة في المدينة والاختلافات الضريبية، بمعنى أنه عند مستوى معنوية 5% يمكننا رفض فرضية العدم بأن $b_1 = 0$ وقبول الفرضية البديلة $b_1 < 0$. نلاحظ، أيضا، أن تقدير b_1 يوحي حجم هذا التأثير يكون عادة كبيرا، فزيادة ١٪ في الاختلافات الضريبية على المبيعات بين المدينة وضواحيها ترتبط بانخفاض قدره ٧٪ في المتوسط، تقريبا، في مبيعات التجزئة في المدينة الرئيسة.*

وينبغي علينا أن ندرك أن هذه النتائج مقنعة إقناعا أكبر من النتائج التي تترتب على الانحدار البسيط لمبيعات المدينة على متغير الضريبة. ومع نتائج

* تمثل المعادلة (4.56) تبسيطاً لنتائج ميكسل. ففي الحقيقة يقترح ميكسل علاقة تأخذ شكل حاصل ضرب على النحو التالي:

$$Y_t = b_0 X_{1t}^{b_1} X_{2t}^{b_2} X_{3t}^{b_3} X_{4t}^{b_4} e^{ut},$$

وبأخذ اللوغاريتمات يصبح لدينا:

$$\ln Y_t = \ln b_0 + b_1 \ln X_{1t} + b_2 \ln X_{2t} + b_3 \ln X_{3t} + b_4 \ln X_{4t} + u_t.$$

ويجدر ذكر أننا سنعالج العلاقات التي تأخذ شكل حاصل ضرب في الفصل القادم، ولكن، نلاحظ أنها تعميم مباشر لمعالجتنا في الفصل السابق للتحويلات اللوغاريتمية.

الانحدار المتعدد هذا، أخذ ميكسل في الحسبان بصراحة تأثير الدخل الفردي ونسبة سكان المدينة إلى سكان الضواحي ومساهمة المدينة على المبيعات. وبمعنى آخر يمكننا القول أن تأثير الضريبة على مبيعات التجزئة بالمدينة قد قيس مع أخذ تأثير المتغيرات الأخرى في الحسبان. وهكذا فإن نتائج ميكسل تتفق مع المقولة بأن الاختلافات في ضريبة المبيعات تسبب تخفيضات مهمة في مبيعات التجزئة في المدن الرئيسة.

ملحق أ (A)

خصائص المقدرات

يهدف هذ الملحق إلى أن يطور، بدقة أكثر، المقدرات وتبايناتها ومعاملات الانحدار، ثم إثبات أن هذه المقدرات غير متحيزة. يضاف إلى ذلك، أن المنهج الذي ستتبعه سيزودنا بمعرفة إضافية حول كيفية قيام تحليل الانحدار المتعدد بفصل تأثيرات مختلف المتغيرات المستقلة.

افترضنا في الفصل الرابع أن بعض المتغيرات المستقلة لنموذج الانحدار ترتبط، في الأقل بالمتغيرات الأخرى كافة أو ببعضها. كما ذكرنا في ذلك الفصل، سنحاول تفسير المتغير المستقل الأخير بدلالة المتغيرات المستقلة الأخرى، عن طريق نموذج الانحدار:

$$X_{kt} = c_0 + c_1 X_{1t} + \cdots + c_{k-1} X_{(k-1)} + v_{kt}. \qquad (4A.1)$$

افترض أننا نقدر معلمات (4A.1) بطريقة المتغير المساعد. وللقيام بذلك سنضع $\Sigma\left(\hat{v}_{kt}X_{(k-1)}t\right)=0$ و ... و $\Sigma\left(\hat{v}_{kt}X_{1t}\right)=0$ و $\Sigma\hat{v}_{kt}=0$ للحصول على مجموعة المعادلات الطبيعية:

$$\begin{aligned}
&\sum X_{kt} = n\hat{c}_0 + \hat{c}_1 \sum X_{1t} + \cdots + \hat{c}_{k-1} \sum X_{(k-1)t} \\
&\sum (X_{kt}X_{1t}) = \hat{c}_0 \sum X_{1t} + \hat{c}_1 \sum X_{1t}^2 + \cdots + \hat{c}_{k-1} \sum (X_{1t}X_{(k-1)t}) \\
&\vdots \\
&\sum \left(X_{kt}X_{(k-1)t}\right) = \hat{c}_0 \sum X_{(k-1)t} + \hat{c}_1 \sum \left(X_{1t}X_{(k-1)t}\right) + \cdots + \hat{c}_{(k-1)} \sum X_{(k-1)t}^2
\end{aligned} \qquad (4A.2)$$

لاحظ وجود معلمات غير معلومة عددها k : $\hat{c}_0$ ، $\hat{c}_1$ ، … ، $\hat{c}_{k-1}$ ومعادلات عددها k في المعادلة (4A.2) لذلك - وبتحقق افتراضاتنا - يمكننا (عموما) أن نحل هذه المعادلات للحصول على تلك المعلمات. ويمكننا هذا من تكوين:

$$\hat{X}_{kt} = \hat{c}_0 + \hat{c}_1 X_{1t} + \cdots + \hat{c}_{k-1} X_{(k-1)t}. \qquad (4A.3)$$

وسيكون مقدر v_{kt} هو:

$$\hat{v}_{kt} = X_{kt} - \hat{X}_{kt}. \qquad (4A.4)$$

وبإعادة ترتيب الحدود في المعادلة (4A.4) يصبح لدينا:

$$X_{kt} = \hat{X}_{kt} - \hat{v}_{kt}. \qquad (4A.5)$$

لاحظ، بدقة، مافعلناه حتى الآن. قسمنا قيمة X_k إلى جزئين: الجزء الأول $\hat{X}_{kt}$ ، وهو ذلك الجزء من X_k الذي يرتبط مباشرة بالمتغيرات المستقلة الأخرى، فهو يمثل التغير في X_k الذي يمكن أن ينسب إلى المتغيرات المستقلة الأخرى X's. وبالمقابل، فإن $\hat{v}_{kt}$ هو ذلك الجزء من X_k الذي لانستطيع نسبته إلى بوساطة المتغيرات المستقلة الأخرى أوتفسيره بوساطتها. لاحظ، مرة أخرى، أنه إذا كان $\hat{v}_{kt}=0$ لجميع قيم t فسوف يكون لدينا ارتباط متعددِ خطي تام [انظر المعادلة (4A.5) و (4A.3)] وحينئذ، لايمكننا تقدير b_k. ولهذا السبب، نتوقع أن يؤدي $\hat{v}_{kt}$ دورا مهما في تقدير b_k.

وبعد أخذ هذا في الحسبان، دعنا نسترجع مماورد في هذا الفصل أنه لدينا عدد (k+1) من المعادلات الطبيعية المحددة لـ$\hat{b}_i$ ، وتأخذ إحداها الشكل:

$$\sum (Y_t X_{kt}) = \hat{b}_0 \sum X_{kt} + \hat{b}_1 \sum (X_{1t} X_{kt}) + \cdots + \hat{b}_k \sum X_{kt}^2. \qquad (4A.6)$$

نعرف من المعادلة (4A.5) أن X_{kt} يمكن التعبير عنها كمجموع كل من $\hat{X}_{kt}$ و $\hat{v}_{kt}$ ، كما نعلم، أيضا، أن $\Sigma\hat{v}_k = 0$ و $\Sigma\hat{v}_{kt}X_{ik} = 0$ حيث k-1 و ...,i = 1,2 . ولذلك طالما أن $\hat{X}_{kt}$ توليفة خطية من X_{1t} ، X_{2k} ، … ، $X_{(k-1)}$ [انظر المعادلة (4A.3)] وأن $\Sigma(\hat{X}_{kt} \ \hat{v}_{kt}) = 0$ فإذا عوضنا عن X_{kt} في كل مكان في المعادلة (4A.6) بـ $(\hat{X}_{kt} + \hat{v}_{kt})$ وألغينا الحدود التي تساوي الصفركافة، فإننا سنحصل على:

$$\sum(Y_t\hat{X}_{kt})+\sum(Y_t\hat{v}_{kt})= \hat{b}_0\sum\hat{X}_{kt}+\hat{b}_1\sum(X_{1t}\hat{X}_{kt})+\cdots+\hat{b}_k\sum\hat{X}_{kt}{}^2+\hat{b}_k\sum\hat{v}_{kt}{}^2. \tag{4A.7}$$

ويتساوى مجموع الحدود الـــ (k+1) الأولى في الجانب الأيمن من المعادلة (4A.7)، ببساطة، مع $\Sigma(Y_t\hat{X}_{kt})$:

$$\sum(Y_t\hat{X}_{kt})=\hat{b}_0\sum\hat{X}_{kt}+\hat{b}_1\sum(X_{1t}\hat{X}_{kt})+\cdots+\hat{b}_k\sum\hat{X}_{kt}{}^2. \tag{4A.8}$$

وحتى نرى ذلك نعوض:

$$Y_t=\hat{b}_0+\hat{b}_1X_{1t}+\cdots+\hat{b}_kX_{kt}+\hat{u}_t$$

في المجموع الأيسر من المعادلة (4A.8) حيث نلاحظ أن:

$$\sum(\hat{u}_t\hat{X}_{kt})=\hat{c}_0\sum\hat{u}_t+\hat{c}_1\sum(\hat{u}_tX_{1t})+\cdots+\hat{c}_{(k-1)}\sum(\hat{u}_tX_{(k-1)t})=0.$$

وسبب كل هذا هو أننا إذا عوضنا المعادلة (4A.8) في المعادلة (4A.7) نحصل على:

$$\sum(Y_t\hat{v}_{kt})=\hat{b}_k\sum\hat{v}_{kt}{}^2, \tag{4A.9}$$

ثم نحل المعادلة (4A.9) للوصول إلى $\hat{b}_k$:

$$\hat{b}_k=\frac{\sum(Y_t\hat{v}_{kt})}{\sum\hat{v}_{kt}{}^2} \tag{4A.10}$$

وباختصار، تعتمد $\hat{b}_k$ فقط على Y's و $\hat{v}_{kt}$ حيث تمثل $\hat{v}_{kt}$ التغير المستقل في X_{kt} وتعميم المعادلة (4A.10) عموما هو:

$$\hat{b}_i=\frac{\sum(Y_t\hat{v}_{it})}{\sum\hat{v}_{it}{}^2} \tag{4A.11}$$

حيث $\hat{v}_{it}$ هو الباقي من انحدار X_{it} على الـ X's الأخرى كافة. أي أن مقدر كل واحد من b_i يمكن التعبير عنه بدلالة قيم Y_t وقيم الحد $\hat{v}_{it}$ الذي يمثل التغير «المستقل»، في المتغير المستقل المقابل، X_{it}.

مقدرات غير متحيزة

في المبحث السابق، أثبتنا أن:

$$\hat{b}_t = \frac{\sum (Y_t \hat{v}_{it})}{\sum \hat{v}_{it}^2} \qquad (4A.11)$$

ولكي نثبت أن $\hat{b}_i$ غير متحيزة، نعيد كتابة النموذج الأساسي في الفصل والموضح في معادلة (4.3) على النحو:

$$Y_t = b_0 + b_1 X_{1t} + \cdots + b_t (\hat{X}_{it} + \hat{v}_{it}) + \cdots + b_k X_{kt} + u_t, \qquad (4A.12)$$

حيث عوضنا عن X_{it} بـ $\left(\hat{X}_{it} + \hat{v}_{it}\right)$. بعد ذلك نضرب المعادلة (4A.12) في $\hat{v}_{it}$ ونجمع على مدى n من المشاهدات، ثم نعوض عن $\Sigma\left(Y_t \hat{v}_{it}\right)$ في المعادلة (4A.11) للحصول على:

$$\hat{b}_i = b_i + \frac{\sum (\hat{v}_{it} \hat{u}_t)}{\sum \hat{v}_{it}^2} \qquad (4A.13)$$

عند حصولنا على المعادلة (4A.13) استخدمنا شروط المعادلة الطبيعية $\Sigma\hat{v}_{it} = 0$ و $\Sigma\left(\hat{v}_{it}\ X_{jt}\right) = 0$ إذا كانت $j \neq i$ ولذا تكون $\Sigma\left(\hat{v}_{it}\ \hat{X}_{it}\right) = 0$. لاحظ أن $\hat{v}_{it}$ لايعتمد على قيم u_t ولكن يعتمد، فقط، على القيم المعطاة لـ X's. وطالما نفترض استقلال u_t عن X's (لأي قيم معطاة لـ X's) فإن $\hat{v}_{it}$ سوف يحصل عليها وسيصبح لدينا:

$$E(\hat{b}_i) = E(b_i) + E\left[\frac{\sum (\hat{v}_{it} u_t)}{\sum \hat{v}_{it}^2}\right]$$

$$= b_i + \left(\frac{\hat{v}_{i1}}{\sum \hat{v}_{it}^2}\right) E(u_1) + \cdots + \left(\frac{\hat{v}_{in}}{\sum \hat{v}_{it}^2}\right) E(u_n) \qquad (4A.14)$$

$$= b_i.$$

وهكذا، فمقدرنا لـ b_i غير متحيز.

تباينات المقدرات

والآن، من السهل علينا أن نشتق التباين الشرطي لـ$\hat{b}_i$ من المعادلة (4A.13) وبالتحديد دعنا نعيد كتابة المعادلة (4A.13) بالتفصيل للحصول على:

$$\hat{b}_i = b_i + \left(\frac{\hat{v}_{i1}}{\Sigma \hat{v}_{it}^2}\right)u_1 + \left(\frac{\hat{v}_{i2}}{\Sigma \hat{v}_{it}^2}\right)u_2 + \cdots + \left(\frac{\hat{v}_{in}}{\Sigma \hat{v}_{it}^2}\right)u_n \qquad (4A.15)$$

وبجعل $M_{it}^2 = \hat{v}_{it} / \Sigma \hat{v}_{it}^2$ يصبح لدينا:

$$\hat{b}_i = b_i + M_{i1}u_1 + \cdots + M_{in}u_n \qquad (4A.16)$$

أي أن $\hat{b}_i$ توليفة خطية من الاخطاء العشوائية. وباستخدام الصيغة الموجودة في الفصل الثاني لتباين المجموع الخطي من المتغيرات العشوائية غير المرتبطة ببعضها بعضا، يكون لدينا:

$$\operatorname{var}(\hat{b}_i) = {M_{i1}}^2\, \sigma_u^2 + {M_{i2}}^2\, \sigma_u^2 + \cdots + {M_{in}}^2\, \sigma_u^2 . \qquad (4A.17)$$

دع $A = \Sigma \hat{v}_{it}^2$. حينئذ يكون $M_{it}^2 = \hat{v}_{it}^2 / A^2$. وباستخدام هذا يمكننا إعادة كتابة المعادلة (4A.17) على النحو:

$$\operatorname{var}(\hat{b}_i) = \frac{\sigma_u^2}{A^2} = ({\hat{v}_{i1}}^2 + {\hat{v}_{i2}}^2 + \cdots + {\hat{v}_{in}}^2)$$

$$= \frac{\sigma_u^2 (\sum {\hat{v}_{it}}^2)}{A^2} \qquad (4A.18)$$

$$= \frac{\sigma_u^2}{A^2} A = \frac{\sigma_u^2}{A}$$

باختصار،

$$\operatorname{var}(\hat{b}_i) = \frac{\sigma_u^2}{\sum {\hat{v}_{it}}^2}, \qquad for \;\; i = 1, \cdots, k \qquad (4A.19)$$

ملحق ب (B): العلاقة بين R^2 و $\overline{R}^2$

أشرنا في هذا الفصل إلــى أن $\overline{R}^2 \geq R^2$. هنا نثبت هذه القاعدة، اعتبــر نموذج الانحدار الخطي التالي:

$$Y_t = b_0 + b_1 X_{1t} + \cdots + b_k X_{kt} + u_t, \tag{4B.1}$$

بحد ثابت وعدد k من المتغيرات المستقلة. لذلك، فإن عدد المعاملات هو $p=k+1$. إذا اعتبرنا النموذج الذي يكون فيه $k \geq 1$، حينئذ، تكــون $p \geq 2$. لمثل هذا النموذج، دع R^2 و $\overline{R}^2$ معامل التحديد ومعامل التحديد المعدل على الترتيب. حينئذ، فإننا سوف نثبت أدناه مايلي:

$$\overline{R}^2 < R^2 \tag{4B.2}$$

إلا إذا كانت $R^2 = 1$ (عندها تكون $\overline{R}^2 = R^2 = 1$). وطالما أن R^2 تكون، عادة، أقل من الواحد الصحيح، فإن النتيجة في المعــادلــة (4B.2) تشير إلى أن $\overline{R}^2$ ستكون، عادة، أقل من R^2.

للحصول على المعادلة (4B.2)، لاحظ من المعادلة (4.52) أن $\overline{R}^2$ يمكن التعبير عنها على النحو:

$$\overline{R}^2 = \left(1 - \frac{ESS}{TSS}\right) + \frac{ESS}{TSS} - \frac{ESS/(n-p)}{TSS/(n-1)}. \tag{4B.3}$$

وبما أن $R^2 = 1 - ESS/TSS$، فإنه ينتج عن ذلك أن $ESS/TSS = 1-R^2$. لاحظ، أيضا، أن:

$$\frac{ESS/(n-p)}{TSS/(n-1)} = \frac{ESS}{TSS}\left(\frac{n-p}{n-1}\right). \tag{4B.4}$$

وينتج عن المعادلة (4B.3) بعدئذ، أن $\overline{R}^2$ يمكن التعبير عنها على النحو:

$$\overline{R}^2 = R^2 + \frac{ESS}{TSS} - \left(\frac{ESS}{TSS}\right)\left(\frac{n-1}{n-p}\right).$$

$$= R^2 + \left(\frac{ESS}{TSS}\right)\left(1 - \frac{n-1}{n-p}\right)$$

$$=R^2+(1-R^2)\left(1-\frac{n-1}{n-p}\right) \tag{4B.5}$$

$$=R^2+(1-R^2)\left(\frac{n-p}{n-p}-\frac{n-1}{n-p}\right)$$

$$=R^2+(1-R^2)\left(\frac{1-p}{n-p}\right)$$

$$=R^2-(1-R^2)\left(\frac{p-1}{n-p}\right)$$

من الواضح الآن أنه، إذا كان $R^2 = 1$ فإن $\overline{R}^2 = R^2$. أما إذا كانت $R^2 < 1$ و $p \geq 2$ فإن $(1-R^2)(p-1)/(n-p) > 0$. وحيئذ، ينتج عن السطر الأخير من المعادلة (4B.5) أن $\overline{R}^2 < R^2$.

أسئلة

١- اعتبر نموذج الانحدار التالي:

$$Y_t = a_0 + a_1X_{1t} + a_2X_{2t} + u_t$$

(ا) أذكر الافتراضات المعتادة لهذا النموذج.

(ب) أذكر المعادلات الطبيعية مشيرا إلى الافتراض الذي يناظر كلا منها.

(ج) افترض أن عينتنا تحتوي على $n = 100$, $\Sigma X_1 = \Sigma X_2 = \Sigma X_1X_2 = 0$ وأن $\Sigma Y = 10$, $\Sigma(YX_1) = 30$, $\Sigma(YX_2) = 20$, $\Sigma X_1^2 = 35$ وأخيرا $\Sigma X_2^2 = 3$. قدر a_0، a_1 و a_2.

٢- اعتبر النموذج:

$$Y_t = a_0 + a_1X_{1t} + a_2X_{2t} + a_3(X_{1t} - X_{2t}) + a_4X_{1t}X_{2t} + \varepsilon_t$$

ماهي المعلمات التي لايمكن تقديرها في ظل افتراضاتنا المعتادة؟ ولماذا؟

٣- اعتبر نموذج الانحدار التالي:

$$Y_t = b_0 + b_1X_{1t} + b_2X_{1t} + u_t$$

حيث إن مشاهدات X_{1t}، X_{2t} و Y_t هي:

X_{1t}	X_{2t}	Y_t
1	2	7
2	1	8
1	3	5
3	1	6
1	2	4

اكتب المعادلات الطبيعية.

٤- يقال إن العائلات عالية الدخل ومتوسطته تنزح من المدن وتعيش في الضواحي بسبب الضرائب العالية نسبيا ومعدلات الجرائم العالية، وتكاليف الاسكان العالية، وأيضا، بسبب أنها ترغب في مكان أوسع للإقامة. كون نموذج انحدار يمكن استخدامه لاختبار مثل هذه الفرضية، اشرح المزايا النسبية للبيانات المقطعية والسلاسل الزمنية لاختبار مثل هذه الفرضيات.

٥- افترض أن: $D_{1t} = a_0 + a_1 p_{1t} + a_2 p_{2t} + \cdots + a_k p_{kt} + b\overline{p}_t + cy_t + u_{1t}$ حيث D_{1t} هو الطلب على السلعة 1، P_{1t}: ثمن السلعة 1، p_{2t}، ...، p_{nt} هي أسعار (k-1) من السلع الأخرى و $\overline{P}_t = \sum_{i=1}^{k} \left(P_{it}\right)/k$ هو المستوى العام للأسعار، وY_t هو الدخل. باختصار، يعتقد بعض الناس أن الطلب على السلعة 1 يعتمد على مستوى سعرها أسعار السلع الأخرى ومستوى الأسعار العام والدخل.

(أ) هل يمكن أن تثار أي مشاكل في تقدير هذا النموذج ؟

(ب) هل يمكن تقدير أي من معلمات النموذج السابق؟ ماهي تلك المعلمات؟ وضح.

٦- اعتبر النموذج: $Y_t = b_0 + b_1 X_t + b_2 X_t^2 + \varepsilon$

(أ) هل المعادلة السابقة تعاني من تعدد العلاقات الخطية التام.

(ب) افترض أن لدينا المشاهدات التالية:

Y	-1	-1	2	4
X	0	1	2	5

اكتب المعادلات الطبيعية.

الفصل الخامس

طرق أخرى في تحليل الانحدار المتعدد

عمّمنا في الفصل السابق طريقة التقدير الأساسية لنموذج انحدار المتغيرين إلى حالة المتغيرات المستقلة المتعددة. في هذا الفصل، سندرس بعض الطرق الإضافية التي يمكن استخدامها في تحليل الانحدار المتعدد. وبالتحديد، سنوسع أولا تحليلنا السابق بمعالجة حالة المتغيرات المبطأة في الانحدار المتعدد، وفي هذا المجال، سنقدم ثلاث طرق لتقدير الاشكال المختلفة للعلاقات المبطأة. سنقدم ثانيا: مفهوم المتغيرات «الصورية» وهي تلك المتغيرات التي تمكننا من حساب بعض التأثيرات النوعية التي تدخل في علاقتنا الاقتصادية. على سبيل المثال، يمكننا استخدام المتغيرات الصورية من الأخذ في الحسبان تأثير عوامل مثل الجنس والدين على أنواع معينة من السلوك. هذه الطريقة تمكننا من أن ندخل في تحليلنا متغيرات لايمكن قياسها قياسا كميا تقليديا. وأخيرا، سنعود إلى مسألة الشكل رقمالدالي وسنرى كيفية استخدام تحليل الانحدار المتعدد لتقدير أنواع مختلفة من العلاقات.

(٥-١) تقدير العلاقات المبطأة

في الفصل الثالث، أوضحنا أن نموذج انحدار المتغيرين يمكن أن يستخدم لتقدير معادلات يعتمد فيها المتغير التابع على قيمة المتغير المستقل في فترة سابقة.

على سبيل المثال، تذكر أننا اعتبرنا الحال، التي يكون فيها الإنفاق الاستهلاكي في فترة معينة يعتمد على الدخل المتاح في الفترة الزمنية السابقة على النحو التالي:

$$C_t = a + bY_{d(t-1)} + u_t. \tag{5.1}$$

ولكن، ليس هناك سبب ضروري لاعتماد الاستهلاك الحالي على الدخل المتاح في الفترة الزمنية السابقة مباشرة، فقط، ذلك أن مستويات الدخل في فترات سابقة، بالإضافة إلى مستواه الحالي قد تؤثر على الإنفاق الاستهلاكي. فإذا كان ذلك صحيحا فسيكون لدينا علاقة من الشكل:

$$C_t = a + b_0Y + b_1Y_{d(t-1)} + \cdots + b_kY_{d(t-k)} + u_t. \tag{5.2}$$

يطلق على هذا النوع من العلاقات فترات الإبطاء الموزعة distributed lag، ويعني هذا أن قيمة المتغير التابع في أي فترة زمنية معطاة تعتمد على مجموع مرجح بالأوزان للقيم الماضية للمتغير المستقل. ويمكننا أن نتخيل، في هذه الحال، أن المتغير التابع C_t سيتكيف ببطء للقيمة الحالية للمتغير المستقل Y_{dt} وذلك بسبب الدفع الذاتي المتراكم من القيم الماضية المبطأة للمتغير المستقل.

ومثال نموذج رياضي لمثل هذه العلاقة بين C_t وY_{dt} هو نموذج تعتمد فيه قيم الإنفاق الاستهلاكي في الفترة t على الدخل المتوقع في الفترة t+1، ولكن، حيث يحدد الدخل المتوقع، بدوره، على أساس مجموع مرجح بالأوزان لمستويات الدخل في الفترات السابقة. افترض، مثلا، أن:

$$C_t = a + bY^e_{d(t+1)} + u_t, \tag{5.3}$$

حيث $Y_{d(t+1)}$ هو الدخل المتوقع في الفترة (t+1). افترض، أيضا، أن الدخل المتوقع هو مجموع مرجح بالأوزان من الدخول الحالية والماضية على النحو:

$$Y^e_{d(t+1)} = \alpha_0Y_{dt} + \alpha_1Y_{d(t-1)} + \cdots + \alpha_kY_{d(t-k)}. \tag{5.4}$$

وبالتعويض من المعادلة (5-4) في المعادلة (5.3) نحصل على:

$$C_t = a + b_0Y_{dt} + b_1Y_{d(t-1)} + \cdots + b_kY_{d(t-k)} + u_t, \tag{5.5}$$

حيث إن $b_0 = b\alpha_0$ و $b_1 = b\alpha_1$ وهلم جرا. وهكذا، سوف تستجيب C_t بتباطؤ

لـY_{dt} بسبب أنY_{dt} هو، فقط، أحد العوامل التي تحدد $Y^{e}_{d(t+1)}$.

وعلى أي حال، يمكن، نظريا، تقدير المعادلة (5.2) مباشرة بطريقة الانحدار المتعدد. افترض أننا نسلم بأن الإنفاق الاستهلاكي الحالي يعتمد على الدخل الحالي وعلى الدخل في السنوات التسع الماضية، لذلك يكون لدينا المعادلة:

$$C_t = a + b_0 Y_{dt} + b_1 Y_{d(t-1)} + \cdots + b_9 Y_{d(t-9)} + u_t . \qquad (5.6)$$

لتقدير المعادلة (5.6)، نحصل على بيانات من النوع الموجود في الجدول رقم (٥-١)، حيث يمثل كل صف مشاهدة مشتركة، ويمكننا استخدام طرق الانحدار المتعدد للحصول على a، b_0، b_1، ...، b_9. أي أن افتراض وجود بيانات من الشكل الموجود في الجدول رقم (٥-١) يمكننا من إعادة تسمية $Y_{d(t-1)}$ أو تعريفها بأنها X_{1t}، $Y_{d(t-2)}$ بأنها X_{2t}، ...، $Y_{d(t-9)}$ بأنها X_{9t} ثم تقدير دالة الاستهلاك (5.6) كما لو أنها معادلة انحدار عادية من الشكل:

$$C_t = a + b_0 Y_{dt} + b_1 Y_{1t} + b_2 X_{2t} + \cdots + b_9 X_{9t} + u_t . \qquad (5.7)$$

وتظهر مشاهداتنا X_{1t}، ...، X_{9t} في الجدول رقم(٥-١).*

جدول رقم (٥-١)

C_t	Y_{dt}	$Y_{d(t-1)}$	$Y_{d(t-2)}$	$\cdots$	$Y_{d(t-9)}$
C_{1950}	$Y_{d(1950)}$	$Y_{d(1949)}$	$Y_{d(1948)}$	$\cdots$	$Y_{d(1941)}$
C_{1951}	$Y_{d(1951)}$	$Y_{d(1950)}$	$Y_{d(1949)}$	$\cdots$	$Y_{d(1942)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
C_{1970}	$Y_{d(1970)}$	$Y_{d(1969)}$	$Y_{d(1968)}$	$\cdots$	$Y_{d(1961)}$

وعلى الرغم من أن هذه العلاقة تعد شكلا ملائما للتقدير في بعض الحالات، إلا أنها تسبب بعض المشاكل. تذكر من الفصل الثالث أنه عندما أدخلنا فترة إبطاء

* يجب على القارئ أن يقنع نفسه، على سبيل المثال، بأن $X_{2(1951)}=Y_{d(1949)}$

واحدة في علاقتنا ذات المتغيرين فقدنا مشاهدة واحدة. ولكن، في حالتنا هـذه، فالوضع أسوأ لأننا نفقد مشاهدة لكل قيمة مبطأة اضافية للدخل المتاح تدخل المعادلة (5.2). افترض، مثلا، أنه يتوافر لدينا قيم مشاهدة للاستهلاك وللدخل المتاح لعشر سنوات من ١٩٦٠م وحتى ١٩٦٩م، وأننا قدرنا المعادلة (5.6) حيث يعتمد الاستهلاك على الدخل الحالي وعلى الدخل في السنوات التسع الماضية. فـي هـذه الحـال، سنفقد تسعا من المشاهدات العشر. أي أن السنة الوحيدة التي تتوافر بها مشاهدات لجميع المتغيرات التي تظهر في المعادلة (5.6) ستكون ١٩٦٩م.

$$C_{1969} = a + b_0 Y_{d(1969)} + b_1 Y_{d(1968)} + \cdots + b_9 Y_{d(1960)} \tag{5.8}$$

وتتطلب علاقة الاستهلاك السابقة لعام ١٩٦٩م استخدام بيانات عن الدخل المتاح قبل ١٩٦٠م غير أنها ليست متاحة. ونتيجة لذلك، فلن تتوافر لدينا مشاهدات كاملة يمكن من خلالها تقدير معادلتنا للانحدار.*

أحد المشاكل الأخرى المرتبطة بالنماذج من النوع (5.6 تنشأ إذا كانت k كبيرة) (يعني ذلك، عـادة، أن $k \geq 5$)، ويعني ذلك وجود عدد كبير من المعلمات التـي ينبغي تقديرها. يضاف إلى ذلك أن هذه المعلمات ستناظر متغيرات ترتبط بصـورة قوية ببعضها البعض. وكما يتوقع القارئ فإن ذلك سيجعل من الصعب عزل تأثير مختلف المتغيرات المستقلة على المتغير التابع (سنبين ذلك في الفصل القادم). بمعنى آخر، ففي ظل هذه الافتراضات، ستكون تباينات مقدرات معلمات الانحدار كبيرة.

---

*ينبغي أن يكون القارئ قادرا، على إثبات أنه إذا كان لدينا مشاهدة واحدة مشتركة فقط، فستكون هناك معادلة طبيعية مستقلة واحدة، لأن جميع المعادلات الطبيعية ستأخذ شكل نسبة من المعادلة الطبيعية الأولى. فمثلا، بالنسبة للمعادلات (5.7) و (5.8) يصبح معامل النسبية في المعادلة الطبيعية الثانية هو وفي المعادلة الطبيعيـة الثالثة هو وهلم جرا [تلميح للحل ستصبح المشاهدة الوحيدة في المعادلة (5.7) أو (5.6) هي عندما يكون t=1، حيث تشير الفترة الزمنية 1 إلى السنة ١٩٦٩م وسيحُصل على المعادلات الطبيعية من الشروط:

$$\sum_{t=1}^{1} \hat{u}_t = \hat{u}_1 = 0, \sum_{t=1}^{1} (\hat{u}_t Y_{dt}) = \hat{u}_1 Y_{dt} = 0, \sum_{t=1}^{1} (\hat{u}_t X_{1t}) = \hat{u}_1 X_{11} = 0, \cdots \sum_{t=1}^{1} (\hat{u}_t X_{9t}) = \hat{u}_1 X_{91} = 0.$$

وهكذا فإن المعادلة الطبيعية الثانية هي Y_{d1} مضروبة في الأولى، والثالثة هي X_{11} مضروبة في الأولى وهلم جرا. تمثل هذه النتيجة حالة خاصة من نتيجة أكثر عمومية تصرح بأنه، ينبغي أن يتوافر لدينا عدد k على الأقل من المشاهدات المشتركة لكي نستطيع تقدير المعلمات.

ثمة مشكلتان أساسيتان مرتبطتان بتحليل فترات الإبطاء الموزعة، الأولى هي المشاهدات التي تفقد بسبب فترات الإبطاء، والثانية هي وجود عدد كبير من المعلمات التي ينبغي تقديرها تقديرا يعتمد عليه. ولمواجهة هذه المشاكل طور الاقتصاديون نماذج لتحليل فترات الإبطاء الموزعة والتي إما أن تقلل من عدد المشاهدات التي تفقد بسبب الإبطاء و / أو تقلل من عدد المعلمات التي ينبغي تقديرها. نعرض الآن لنوعين من هذه النماذج.

إبطاء كويك The Koyck Lag

النموذج الأول هو نموذج، إبطاء كويك*. يستخدم هذا النموذج افتراضا يرتبط بمعلمات العلاقة ذات فترات الإبطاء الموزعة، ويسمح بترجمة هذه العلاقة إلى شكل أبسط كثيرا. وينتج عن هذا النموذج فترات إبطاء أقل وعددا أقل من المعلمات التي يتم تقديرها. ولكن، لسوء الحظ، فإن هذا الشكل رقمالابسط من العلاقة ينتج عنه تعقيدات خطيرة كثيرا مايتم تجاهلها. ولأن نموذج كويك قد شاع استخدامه، فإننا نعرض له هنا ونبين أوجه قصوره، يضاف إلى ذلك أن عرضه سيخدمنا باعتباره خطوة مهمة في المناقشات التالية.

افترض أنه، على الرغم من أن الاستهلاك يعتمد على الدخل في السنوات السابقة فإن تأثير الدخل في السنوات الماضية البعيدة أقل من تأثير الدخل في السنوات الأخيرة. وبالتحديد، افترض أن الاستهلاك الحالي هو مجموع مرجح بالأوزان من السنوات الحالية والماضية للدخل المتاح (مضافا إليه الخطأ العشوائي) وأن هذه الأوزان تتناقص متتاليا للفترات الأكثر بعدا في الماضي. يفترض نموذج كويك تناقص هذه الأوزان هندسيا. على سبيل المثال، اجعل λ ثابتا تتراوح قيمته بين الصفر والواحد الصحيح. حينئذ، وبالرجوع إلى المعادلة (5.2)، يأخذ نموذج كويك الشكل:

* يرجع إلى: .L. M. Koyck. *Distributed Lags and Investment Analysis* (Amsterdam: North Holland, 1954)

$$b_i = \lambda^i b_0, \qquad i = 1,2,\cdots,k. \tag{5.9}$$

وبالتعويض من المعادلة (5.9) في المعادلة (56.2)، يصبح لدينا:

$$C_t = a + b_0 Y_{dt} + (b_0\lambda) Y_{d(t-1)} + (b_0\lambda^2) Y_{d(t-2)} + \cdots + b_0\lambda^k) Y_{d(t-k)} + u_t. \tag{5.10}$$

تبين المعادلـة (5.10) أن الاستهلاك يعتمد على كل من مستويات الدخل الحـالـيـة والماضية، ولكن طالما أن λ مرفوعة إلى قوى أعلى فإنها تصغر باستمرار، وهكذا، تصغر معاملات السنوات الأقدم بالتوالي كلما توغلنا في الماضي.

سيتضح لنا الآن متضمنات نموذج الإبطاء لكويك. فإذا قمنا بإبطاء المعادلـة (5.10)لفترة زمنية واحدة، ثم ضربنا المعادلة الناتجة في λ نحصل على:

$$\begin{aligned}\lambda C_{t-1} &= \lambda a + (\lambda b_0) Y_{d(t-1)} \\ &\quad + (\lambda^2 b_0) Y_{d(t-2)} + \cdots + (\lambda^{k+1} b_0) Y_{d(t-k-1)} + \lambda u_{t-1}.\end{aligned} \tag{5.11}$$

وبطرح المعادلة (5.11) من المعادلة (5.10)، نحصل على:

$$C_t - \lambda C_{t-1} = (a - \lambda a) + b_0 Y_{dt} - \lambda^{k+1} b_0 Y_{d(t-k-1)} + (u_t - \lambda u_{t-1}) \tag{5.12}$$

وبإعادة ترتيب جدول (5.12) يصبح لدينا:

$$C_t = (a - \lambda a) + b_0 Y_{dt} + \lambda C_{t-1} - (\lambda^{k+1} b_0) Y_{d(t-k-1)} + (u_t - \lambda u_{t-1}). \tag{5.13}$$

وبفرض أن k كبيرة (بمعنى وجود عدد كبير من سنوات الإبطاء في الـنـمـوذج)، فسيكون الحد قبل الأخير في المعادلة (5.13) (أو $(\lambda^{k+1} b_0) Y_{d(t-k-1)}$) صغيرا. ومن ثم، وعلى سبيل التقريب، سنضع هذا الحد مساويا الصـفـر*. لتبسيط الرمـوز أيضا، دع $a^* = (a - \lambda a)$ حيث يمكن تبسيط المعادلة (5.13) إلى:

$$C_t = a^* + b_0 Y_{dt} + \lambda C_{t-1} + (u_t - \lambda u_{t-1}). \tag{5.14}$$

والآن دع:

$$v_t = (u_t - \lambda u_{t-1}). \tag{5.15}$$

---

* وبالطبع إذا افترضنـا أن k تؤول إلى مالانهاية بينما يظل الحد $Y_{d(t-d-1)}$ محدودا فإن هذا الحد سيسـاوي الصفر. هذا هو الافتراض المعتاد الذي يستخدم في الدراسات القياسية.

طالما أن v_t يعتمد، فقط، على الأخطاء العشوائية، فمن المنطقي اعتبار v_t نفسه خطأ عشوائيا، دعنا للحظة، نفترض أن v_t (والذي، لسوء الحظ، غالبا مايفترض في الدراسات القياسية) يحقق افتراضات نموذج الانحدار المرتبطة بخصائص الخطأ العشوائي كافة. ويمكن التعبير في هذه الحال، عن المعادلة (5.14) على النحو التالي:

$$C_t = a^* + b_0 Y_{dt} + \lambda C_{t-1} + v_t , \tag{5.16}$$

حيث a ثابت وأن:

$$E(v_t) = 0,$$
$$E(v_t^2) = \sigma_v^2 ,$$
$$E(v_t v_{t-1}) = 0,$$
$$E(v_t Y_{dt}) = E(v_t C_{t-1}) = 0.$$

لاحظ حجم التبسيط الهائل الذي تمثله المعادلة (5.16) مقارنة بالمعادلة (5.2) حيث جمع تأثير القيم المبطأة كافة للدخل المتاح على الاستهلاك الحالي في حد واحد: قيمة الاستهلاك المبطأة لفترة زمنية واحدة. نحتاج الآن، فقط، لتقدير قيمة λ بدلا من معاملات القيم المبطأة كافة، وبمعنى آخر إذا قبلنا نموذج كويك لفترات الإبطاء، وقبلنا، أيضا، الافتراضات المرتبطة بـv_t، فإن معلمات نموذج مثل (5.2) يمكن تقديرها بدلالة نموذج مثل (5.16)، والذي يتطلب تقديره خسارة مشاهدة واحدة فقط.

وحتى نكون أكثر تحديدا لطريقة التقدير، يمكن اعتبار المعادلة (5.16) نموذجا للانحدار المتعدد يتضمن المعلمات a^*، b وأخيرا λ. وسنحصل على المقدرات $\hat{a}^*$، $\hat{b}_0$ وأخيرا $\hat{\lambda}$ وذلك بالتطبيق المباشر لطريقة المتغير المساعد. وباستخدام هذه المعلمات المقدرة نحصل على مقدرات، المعلمات كافة في المعادلة (5.2). فعلى سبيل المثال، إذا افترضنا أنه يوجد عدد لانهائي من فترات الإبطاء (أي k لانهائية):

$$\hat{b}_i = (\hat{\lambda}^i)\hat{b}_0, \qquad i = 1,2,\cdots, \tag{5.17}$$

وحيث إن $a^* = a - \lambda a$ ، ولذلك فإن مقدر a سيكون:

$$\hat{a} = \frac{\hat{a}^*}{1-\hat{\lambda}} \tag{5.18}$$

وللتعبير عن النتائج بعمومية أكثر، تمكننا افتراضات نموذج كويك (مشتملة على افتراض أن k لانهائية) من وضع النموذج:

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \cdots + u_t \tag{5.19}$$

في الشكل المبسط:

$$Y_t = a^* + b_0 X_t + \lambda Y_{t-1} + v_t , \tag{5.20}$$

حيث نحتاج، فقط، لتقدير ثلاث معلمات a^*، b_0 و λ. وتصبح العلاقة بين معلمات المعادلة (5.19) والمعادلة (5.20) هي:

$$a = \frac{a^*}{1-\lambda}, \qquad b_i = \lambda^i b_0 , \qquad i = 1,2,\cdots \tag{5.21}$$

وهكذا، فعن طريق تحويل النموذج صراحة إلى نموذج يحتوي على فترة إبطاء، واحدة، يؤدي استخدام نموذج كويك لفقد مشاهدة مشتركة واحدة. ومزية نموذج كويك ينبغي أن تكون الآن واضحة.

ولكن - وكما لاحظنا من قبل - هناك بعض المشاكل الصعبة مع نموذج كويك. تذكر أولا أننا عند اشتقاق المعادلة التي تقدرها، افترضنا أن الخطأ العشوائي

$$v_t = (u_t - \lambda u_{t-1})$$

يحقق الشروط كافة التي تفرضها عادة الأخطاء العشوائية. ولكن، لسوء الحظ، فإن ذلك ليس صحيحا بالضرورة. فإذا كانت u_t في المعادلة الأصلية (5.19) تحقق افتراضات نموذج الانحدار كافة، فإن v_t في المعادلة (5.20) لاتحققها، عمومًا. وبالتحديد، فإن قيم v_t سيكون لها ارتباط غير صفري مع بعضها بعضا. وأيضا، مع واحد من المتغيرات المستقلة -القيمة المبطأة للمتغير التابع Y_t. على سبيل المثال، إذا كانت $v_t = u_t - \lambda u_{t-1}$ و $v_{t-1} = u_{t-1} - \lambda u_{t-2}$ حينئذ، فإن v_t و v_{t-1} لن يكونا مستقلين عن بعضهما بعضا طالما أنهما يحتويان على حد مشترك (u_{t-1}). ومن ثم، لانتوقع أن تتحقق $E(v_t\, v_{t-1}) = 0$. وفي الحقيقة، فإنه، في ظل افتراضاتنا العادية المتعلقة بـ u_t، لدينا:

$$
\begin{aligned}
E(v_t v_{t-1}) &= E[(u_t - \lambda u_{t-1})(u_{t-1} - \lambda u_{t-2})] \\
&= E(u_t u_{t-1} - u_t \lambda u_{t-2} - \lambda u_{t-1}^2 + \lambda^2 u_{t-1} u_{t-2}) \qquad (5.22) \\
&= -\lambda \sigma_u^2 \neq 0.
\end{aligned}
$$

وهكذا فإن التغاير بين القيم المتعاقبة للأخطاء العشوائية في المعادلة (5.20) لن يساوي الصفر. ونترك للقارئ، على سبيل التدريب، أن يثبت بطريقة مشابهة أن*:

$$E(v_t Y_{t-1}) \neq 0. \qquad (5.23)$$

وباختصار، إذا بدأنا بمعادلة من الشكل رقم(5.19) والتي تحقق افتراضات نموذج الانحدار كافة فإن تحويل كويك لهذه المعادلة سيؤدي، عمومًا إلى انتهاك بعض هذه الافتراضات.** يضاف إلى ذلك أن النتائج المترتبة على هذه الانتهاكات خطيرة. فمثلا – وكما سنوضح في فصل لاحق – تتضمن المعادلة (5.23) أن مقدراتb_i, b و λ لن تكون متحيزة وحسب بل غير متسقة ايضا.***

بالإضافة إلى مشاكل التقدير التي سبق الإشارة إليها، يعد نموذج كويك مفيدا جدا من حيث إنه يفترض تناقص تأثير الفترات الماضية تأثيرا متعاقبا وبطريقة معينة. وهذا، بالتأكيد، لن يكون صحيحا دائما. فعلى سبيل المثال، – وبسبب تأثير العادة – فقد يكون صحيحا أن مستوى الدخل في الفترة السابقة مباشرة أكثر أهمية في التأثير على مستوى الاستهلاك الحالي عن مستوى الدخل الحالي. ومثل هذه العلاقة لاتتناسق بوضوح مع نموذج كويك لفترات الإبطاء. وهكذا يصبح من المفيد جدا الحصول على نموذج لفترات الإبطاء الموزعة يتصف بمرونة أكبر من نموذج كويك، كما لايخالف افتراضات نموذج الانحدار.

---

* تلميح للحل: بما أن Y_t من المعادلة (5.20) تعتمد مباشرة على v_t، فإن Y_{t-1} تعتمد على V_{t-1} : $Y_{t-1} = a^* + b_0 X_{t-1} + \lambda Y_{t-2} + V_{t-1}$ ولذلك سترتبط Y_{t-1} بداهة بحكم أن v_t و Y_{t-1} غير مستقلين عن بعضهما بعضا. لإثبات المعادلة (5.23)، اضرب Y_{t-1} في v_t ثم خذ التوقعات. وعند القيام بذلك لاحظ أن $E[Y_{t-2} \; V_t] = 0$.

** سنعرض الطرق التي تستخدم لمواجهة انتهاكات افتراضات نموذج الانحدار في الفصلين السادس والسابع.

*** لأن المعادلة الطبيعية الثالثة من نموذج الانحدار (5.20) مبنية على الشرط $\Sigma(\hat{v}_t Y_{t-1}) = 0$، الذي لم يعد متسقاً مع المعادلة (5.23). سنعرض مزيدًا من التوضيح فيما بعد.

إبطاء آلمون *The Almon Lag

على الرغم من أن المشاكل التي تظهر في نموذج كويك يمكن مواجهتها في اطار نموذج أكثر عمومية سنكونه فيما بعد، فإن الحل ليس سهلا. ويمكن التغلب على هذه الصعوبة عن طريق استخدام نموذج آلمون. ونؤكد هنا أن هذا النموذج لايقلل - كما هو الحال في طريقة كويك - عدد المشاهدات التي تفقد بسبب وجود المتغيرات المبطأة، ولكنه يقلل، فعلا، عدد المعلمات التي تقدر. إضافة إلى ذلك، هناك ميزتان لنموذج آلمون على نموذج كويك. الأولى أنه لايخالف أي من افتراضات نموذج الانحدار. والثانية - وكما سنرى - أنه أكثر مرونة بدرجة كبيرة من نموذج كويك بدلالة أشكال الإبطاء وهياكله المسموح بها.

بالعودة إلى التكوين العام للعلاقة المبطأة، دعنا نفترض أن النموذج الذي نرغب في تقريره يأخذ - مرة أخرى - الشكل:

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \cdots + b_k X_{t-k} + u_t, \tag{5.24}$$

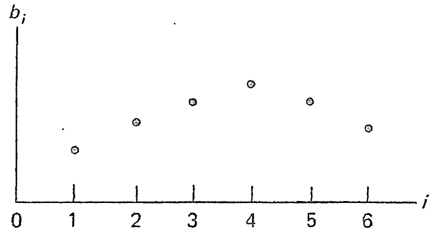
حيث يستوفي الخطأ العشوائي u_t الافتراضات المعتادة. وكما سبق، قد نتوقع أن تكون معاملات الـ X's المناظرة لفترات زمنية بعيدة أقل من تلك المعاملات المناظرة للفترات الزمنية الأقرب. ولكن، من ناحية أخرى، ولاعتبارات معينة، فإن ذلك لن يكون صحيحا بالضرورة في نموذج (5.24)، ففي بعض الحالات تزيد b's فعلا في البداية (أي $b_1 > b_0$) ثم تبدأ بعد ذلك في التناقص تدريجيا. فمثلا، بسبب البطء في الإدراك، يحدث تباطؤ زمني في جمع البيانات، أو في عنصر الزمن المتضمن في اتخاذ القرارات، فالإنفاق الاستثماري، مثلا، قد يكون أكثر استجابة لظروف الطلب في فترات زمنية قليلة سابقة عن ظروف الطلب الحالية. وباختصار، وفي ظل افتراضات مختلفة، يمكن أن نتوقع أنماطا مختلفة لقيم b's في نموذج مثل (5.24).

---

* انظر:

Shirley Almon. "The Distributed Lag between Capital Approapriations and Expenditures." *Econometrica*, 33 (Jan. 1965), pp. 178-196, and Ray Fair and Dwight Jaffee. " A Note on the Estimation of Polynomial Distributed Lags", *Econometrica Research Memorandum*, No. 120, Princeton University, Feb. 1971.

وعلى العكس من طريقة كويك، لاتفترض طريقة آلمون مثل هذه العلاقات الجامدة بين b's. بدلا من ذلك تفترض أنه أيا كان نمط الـ b's فإن ذلك النمط يمكن تحديده بمتعدد للحدود (polynomial). وعلى سبيل المثال، إذا توقعنا أن b's تزداد في البداية ثم تتناقص بعد ذلك، فإن هذا النمط قد يشبه ذلك الموجود في الشكل رقم(٥-١).



شكل رقم (٥-١)

افترض أنه يمكننا تكوين منحنى متصل يمر بالنقاط في الشكل رقم (٥-١) كما فعلنا في الشكل رقم(٥-٢). نرغب الآن في وصف المنحنى بالشكل رقم(٥-٢) جبريا. في الرياضيات توجد نظرية تصرح بأنه-وفي ظل شروط عامة- يمكن تقريب منحنى بوساطة متعدد الحدود. والقاعدة لتحديد درجة متعدد الحدود* هي أن تكون تلك الدرجة المختارة أكبر بمقدار واحد، في الأقل، من عدد نقاط الانقلاب Infletion Points في المنحنى. وبتطبيق هذه القاعدة على الشكل رقم(٥-٢) يمكننا تقريب المنحنى بوساطة متعدد الحدود من الدرجة الثانية، وبالتحديد وباستخدام طريقة آلمون، سنفترض:

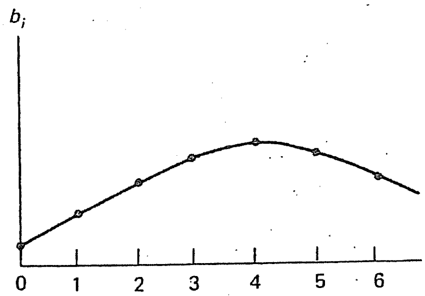
* تذكر من علم الجبر أن «درجة متعدد الحدود» تشير إلى أعلى قوة يرفع إليها المتغير. وهكذا فإن:

$$Y = b_0 + b_1 X + b_1 X^2$$

هو متعدد للحدود من الدرجة الثانية بينما:

$$Y = b_0 + b_1 X + b_2 X^2 + b_3 X^3$$

هو متعدد للحدود من الدرجة الثالثة.



شكل رقم (٥-٢)

$$b_i = a_0 + a_1 i + a_2 i^2, \qquad (5.25)$$

حيث إن α_0، α_1 و α_2 ثوابت ينبغي تحديدها. لاحظ أنه، إذا كانت المعادلة (5.25) هي تقريب للمنحنى في الشكل رقم(٥-٢) فإنه ينبغي أن يكون لدينا:

$$\begin{aligned}
b_0 &= \alpha_0 && (set\ \ i=0) \\
b_1 &= \alpha_0 + \alpha_1 + \alpha_2 && (set\ \ i=1) \\
b_2 &= \alpha_0 + 2\alpha_1 + 2\alpha_2 && (set\ \ i=2) \\
&\vdots \\
b_k &= \alpha_0 + k\alpha_1 + k^2\alpha_2 && (set\ \ i=k)
\end{aligned} \qquad (5.26)$$

وتشتق صيغة كل من b_i في المعادلة (5.26) مباشرة وببساطة من المعادلة (5.25) عن طريق وضع i مساوية قيمة الدليل السفلي لمعامل معين.

قد تبدو المعادلة (5.25) غريبة بعض الشئ عند النظر إليها من حيث إن قيمة الوزن المبطئة b_i يرتبط بطول الفجوة الزمنية ذاتها i. وفي الحقيقة، فإننا واجهنا علاقة مشابهة في طريقة كويك. فهناك افترضنا أن:

$$b_i = \lambda^i b_0. \qquad (5.27)$$

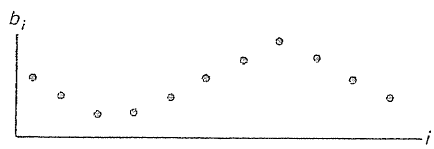
في هذه المعادلة، يرتبط b_i مرة أخرى بـــ i. والاختلاف الوحيد بين المعادلة (5.27) والمعادلة (5.25) هو شكل المعادلة.

قبل أن ننفذ طريقة آلمون، دعنا نوضح، باختصار، إلى أي مدى تتسم هذه الطريقة بالمرونة. افترض أننا نشعر بأن b's تتبع نمطا مثل ذلك الموجود في الشكل رقم(٥-٣) حينئذ سنفترض، ببساطة، أن:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \alpha_3 i^3, \tag{5.28}$$

التي تتضمن أن:

$$\begin{aligned} b_0 &= \alpha_0, \\ b_1 &= \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3, \\ b_2 &= \alpha_0 + 2\alpha_1 + 4\alpha_2 + 8\alpha_3, \\ &\vdots \\ b_k &= \alpha_0 + k\alpha_1 + k^2\alpha_2 + k^3\alpha_3. \end{aligned} \tag{5.29}$$



شكل رقم (٥-٣)

وبشمولية أكثر، لاستخدام طريقة آلمون، كل ماعلينا فعله هو حساب عدد نقاط الانقلاب في نمطنا المفترض لـ b's، ثم نعبر عن b's بوصفه متعددا للحدود في i، حيث تكون درجة متعدد الحدود واحد مضاف إليها عدد نقاط الانقلاب في المنحنى. في الشكل رقم(٥-٤) رسمنا عددا من أنماط فترات الإبطاء الممكنة مع درجة متعدد الحدود المناظر.

سنبين الآن كيفية استخدام طريقة آلمون لتقدير العلاقة التي تنطوي على فترة إبطاء. دعنا الآن نعود إلى التكوين العام الذي اعتبرناه سابقا:

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \cdots + b_k X_{t-k} + u_t, \tag{5.24}$$

افترض أن النظرية الاقتصادية توحي أن متعدد الحدود من الدرجة الثانية مناسب لتحديد شكل فترة الإبطاء. سنأخذ في هذه الحالة:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2. \tag{5.30}$$

فإذا عوضنا عن b's في المعادلة (5.24) بوساطة صيغها الموجودة في المعادلة (5.30) نحصل على:

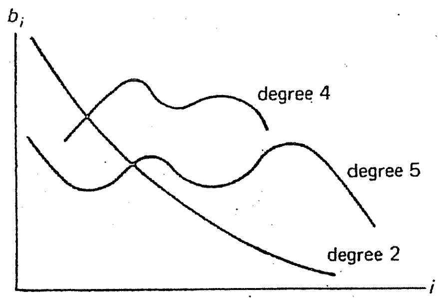
$$\begin{aligned} Y_t = {} & a + \alpha_0 X_t + (\alpha_0 + \alpha_1 + \alpha_2) X_{t-1} \\ & + (\alpha_0 + 2\alpha_1 + 4\alpha_2) X_{t-2} + \cdots + (\alpha_0 + k\alpha_1 + k^2\alpha_2) X_{t-k} + u_t. \end{aligned} \tag{5.31}$$

وبإعادة ترتيب الحدود في المعادلة (5.31)، نحصل على:

$$Y_t = a + \alpha_0 \left(\sum_{i=0}^{k} X_{t-1} \right) + \alpha_1 \left(\sum_{i=1}^{k} i X_{t-i} \right) + \alpha_2 \left(\sum_{i=1}^{k} i^2 X_{t-i} \right) + u_t. \tag{5.32}$$

دعنا الآن نبسط رموزنا عن طريق تعريف:

$$Z_{1t} = \sum_{i=0}^{k} X_{t-i}, \quad Z_{2t} = \sum_{i=1}^{k} i X_{t-i}, \quad and \quad Z_{3t} = \sum_{i=t}^{k} i^2 X_{t-i}. \tag{5.33}$$



شكل (٥-٤)

حيث يمكننا إعادة كتابة المعادلة (5.32) على النحو:

$$Y_t = a + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + u_t . \tag{5.34}$$

المعادلة (5.34) هي نموذج انحدار متعدد عادي يربط Y_t بـ Z_{1t}، Z_{2t} و Z_{3t}. وبسهولة، يمكن أن نوجد مقدرات لكل من a، α_0، α_1 وα_2 بطريقة التقدير العادية. دع $\hat{\alpha}$، $\hat{\alpha}_0$، $\hat{\alpha}_1$ و$\hat{\alpha}_2$ هي هذه المقدرات. حينئذ، يمكننا أن نرى من المعادلة (5.30) أن مقدراتنا لـ b's هي:

$$\begin{aligned}
\hat{b}_0 &= \hat{\alpha}_0 ,\\
\hat{b}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 \\
b_2 &= \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2 \\
&\vdots \\
b_k &= \hat{\alpha}_0 + k\hat{\alpha}_1 + k^2\hat{\alpha}_2 .
\end{aligned} \tag{5.35}$$

لاحظ أنه، بهذه الطريقة، أمكننا الحصول على مقدرات لعدد k من المعلمات b_0، ...، b_k عن طريق الحصول على مقدرات لهذه المعلمات الثلاثة α_0، α_1 وأخيرا α_2. والآن يمكن إثبات أنه، في أي حالة يكون فيها عدد أكبر من b's عن الـ α's تكون مقدرات b's التي حُصل عليها من طريقة آلمون أفضل (بمعنى أنها تحتوي على تباينات أصغر) من المقدرات المباشرة لـ b's التي حُصل عليها بوساطة تطبيق طريقة الانحدار المتعدد مباشرة على المعادلة (5.24).* ولسوء الحظ فإننا لانستطيع اعطاء صيغة مبسطة لتباين المقدرات b_0, ..., b_k التي يحُصل عليها بوساطة طريقة آلمون. ولكن في التطبيق سيزودنا الحاسوب بتقديرات لهذه التباينات حتى نقوم بالاختبارات الاحصائية المعتادة المرتبطة بقيم معاملات الانحدار.

إن تعميم طريقة آلمون واستخدام صيغ مختلفة منها سهل ومباشر. افترض،

* كما نتوقع، فإن هذه النتيجة تكون مبنية على افتراض أن العلاقة المفترضة بين الـ b's والـ α's مثل المعادلة (5.30) صحيحة.

مثلا، أن المعادلة (5.24) قد وسعت لتشتمل على متغير آخر:

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \cdots + b_k X_{t-k} + cW_t + u_t , \qquad (5.36)$$

حيث W_t هو متغير مستقل آخر. إذا افترضنا مـرة أخـرى أن b's تحدد بعلاقة مثل (5.30)فإنه يمكننا أن نتبع الخطوات نفسها لاختزال المعادلة (5.36) لمعادلة تماثل المعادلة (5.24) مع استثناء وحيد وهو وجود الحدcW_t:

$$Y_t = a + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + cW_t + u_t . \qquad (5.37)$$

أي أن ادخال المتغيرات الإضافية لايؤثر أي تأثير على تحليلنا. وفي الحقيـقـة، يمكننا أن نطبق طريقة آلمون لكل من المتغيرات المستقلة المبطئة العديدة في المعادلة نفسها.

ينبغي علينا تذكر ملاحظتين إضافيتين مرتبطتين باستخدام منهج الإبطاء لآلمون، الأولى هي أن المستخدم قد يرغب في وضع بعض القيود الطرفية endpoints على قيم b's فقد نرغب في جعل أما أن b_0 أو b_k (أو كليهما) تعادل الصفر. وبسبب التأخير في تلقي المعلومات، فقد نعتقد، مثلا، أن قيمة المتغير المستقل في الفترة الجـاريـة لاتؤثر على السلوك الحالي (أي على قيمة المتغير التابع في هذه الفترة)، أي أنه في المعادلة (5.34)، نجعل $b_0 = 0$. ويعني هذا أن المتغير التابع (Y) سيعتمد على القيـم المبطئة لـ X في المعادلـة (5.24). ومن الناحية الأخرى، وربما بسبب الطريقة التـي تتخذ بها القرارات، قد نعتقد أن قيم X المبطأة k أو أكثر من الفترات لاتأثير لها على Y، لذلك، نجعل$b_k = 0$ في المعادلة (5.24).

إن إحدى طرق تضمين النموذج مثل$b_k = 0$ هي، ببساطة، إسقاط المتغير X_{t-k} من المعادلة الأساسيـة (5.24)، وإكمال التحليل كماسبق. ولكن، ليست هـذه هـي الطريقة التي تتبع، عادة. فعلى الصعيد العملي، تترجم المعلومات التي تبين أن$b_0 = 0$ أو$b_k = 0$ أو كليهما، باستخدام الافتراض الأساسي مثل المعـادلـة (5.30) شرطا لـ $\alpha's$، ثم تقدير المعادلة الناتجة حينئذ. وعلى الرغم من أن إثبات ذلك يتجاوز مجال هذا الكتاب، فإن هذا المنهج غير المباشر يستخدم لأنه في ظل تحقق بعض الافتراضات يصبح تباين المقدرات غير المتحيزة الناتج أصغر من نظيره إذا اتبع المنهج المـبـاشـر المتلخص في اسقاطX_t و X_{t-k}. وعلى القارئ المهتم أن ينظر إلى الملحق أ (A) لهذا

الفصل حيث تعالج هذه المسألة بتفصيل أكبر، ويوضح الملحق كيف يمكن لمنهج آلمون تضمين هذه الشروط، ولكنه يقترح، أيضا، أسبابا لإتباع المنهج المباشر في التطبيق. ولهذا بعض الأهمية لأن معظم برامج الحاسوب لآلمون تتطلب من المستخدم تحديد القيود الطرفية في حال وجودها.

الثانية: لابد أنك قد لاحظت أننا قد عرضنا لمادة هذا الجزء كما لو أن الباحث يعرف كلا من فترة الإبطاء في نموذجه للانحدار (أي قيمة k) والنمط العام لـ b's وذلك لتحديد درجة متعدد الحدود. إلا أنه، في التطبيق، قد لانعرف أيا من k أو نمط الـ b's. في مثل هذه الحال، ننصح باتباع المنهج التالي: أولا، حدد درجة لمتعدد للحدود (مثلا d) عالية بدرجة كافية حتي تلائم أي نمط معقول لـ b's. في معظم الحالات الدرجة الثالثة أو الرابعة لمتعدد الحدود تكون كافية. افترض أن فترة الإبطاء القصوى «المعقولة» التي نعتقد باتساقها مع العلاقة موضع الاهتمام هي k^*. وعلى سبيل المثال، إذا تم استخدام بيانات ربع سنوية، فإن k^* قد تكون 12 أو 16 والتي تناظر 3 إلى 4 سنوات فترات إبطاء. أما إذا استخدمت بيانات شهرية فربما نأخذ k^* لتعادل 36. وعلى أي حال فعند اختيارك لـ d و k^* قدر العلاقة موضع الاهتمام عند $(k = d, d + 1, ..., k^*)$ حيث d هي درجة متعدد الحدود. لاحظ أننا نهتم فقط بفترات الإبطاء حيث $k \geq d$ لأننا نفترض أن طول فترة الإبطاء تكون على الأقل بطول d (هناك على الأقل عدد من الـ b's بقدر عدد الـ α's).* ينبغي أن تقدر كافة معادلات الانحدار التي تناظر القيم المختلفة لـ k بالبيانات نفسها. لاحظ أن هذا يتطلب أن نهمل المشاهدات الـ k^* الأولى وأن تستخدم فقط الـ $(n\text{-}k^*)$ المتبقية من المشاهدات لتقدير معادلات الانحدار. وسيسمح لنا هذا بمقارنة احصائية R^2 لمختلف المعادلات طالما أنها تؤسس على العينة نفسها. وهكذا ينبغي أن تأخذ قيمة k التي تعظم إحصائية R^2. وفي ظل افتراضات معينة يمكن إثبات أن هذا المنهج يؤدي إلى مقدرات متسقة لكل من k ومعاملات الانحدار.

---

* تذكر أن هدف طريقة آلمون للإبطاء هو تقليل عدد المعلمات التي ينبغي تقديرها. وهذا لن يحدث إذا كانت $k < d$.

مثال

لشرح طريقة آلمون في التقدير ببساطة نعود إلى دالتنا السابقة للاستهلاك. في الجدول رقم (٥-٢) نعيد كتابة المشاهدات العشر حول الإنفاق الاستهلاكي والدخل المتاح للسنوات ١٩٦٠-١٩٦٩م والتي استخدمناها لتقدير معادلتنا التوضيحية للاستهلاك في الفصل الثاني (انظر الجدول رقم ٢-٢). سنستخدم هذه البيانات مرة أخرى لأغراض التقدير. ولكن، في حالتنا هذه، نفترض أن الإنفاق الاستهلاكي يعتمد على الدخل المتاح ذا فترات الإبطاء الموزعة. وأكثر تحديدا دعنا نفترض أن الاستهلاك يعتمد على الدخل المتاح في السنة الحالية وفي السنوات الأربع السابقة. إضافة إلى ذلك، نفترض أن فترة الإبطاء تأخذ شكل متعدد الحدود من الدرجة الثانية. وهكذا تأخذ دالتنا للاستهلاك الشكل رقم التالي:

$$C_t = a + b_0 Y_{dt} + b_1 Y_{d(t-1)} + b_2 Y_{d(t-2)} + b_3 Y_{d(t-3)} + b_4 Y_{d(t-4)} + u_t, \qquad (5.38)$$

حيث:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2.$$

جدول رقم (٥-٢) الاستهلاك والدخل المتاح في الولايات المتحدة ببلايين الدولارات الجارية

السنة	الاستهلاك (C)	الدخل المتاح Y_d
١٩٦٠	٣٢٥	٣٥٠
١٩٦١	٣٣٥	٣٦٤
١٩٦٢	٣٥٥	٣٨٥
١٩٦٣	٣٧٥	٤٠٥
١٩٦٤	٤٠١	٤٣٨
١٩٦٥	٤٣٣	٤٧٣
١٩٦٦	٤٦٦	٥١٢
١٩٦٧	٤٩٢	٥٤٧
١٩٦٨	٥٣٧	٥٩٠
١٩٦٩	٥٧٦	٦٣٠

المصدر: التقرير الاقتصادي للرئيس، واشنطن .D.C، مكتب الطباعة الحكومي، فبراير ١٩٧٠، الصفحات ١٨٩ و ١٩٥.

* تذكر أن هدف طريقة آلمون للإبطاء هو تقليل عدد المعلمات التي ينبغي تقديرها. وهذا لن يحدث إذا كانت k < d.

ولوضع معادلتنا للاستهلاك في شكل المون ينبغي أن نحسب:

$$Z_{1t} = \sum_{i=0}^{4} Y_{d(t-i)}, \qquad Z_{2t} = \sum_{i=1}^{4} i\, Y_{d(t-i)}, \qquad \text{and} \qquad Z_{3t} = \sum_{i=1}^{4} i^2\, Y_{d(t-1)}.$$

وتظهر قيم Z's مع قيم المتغير التابع في الجدول رقم (٥-٣). لاحظ أننا نتيجة وجود فترة إبطاء لأربع سنوات نفقد أربعا من مشاهداتنا، ولذلك فالجدول رقم (٥-٣) يحتوي على بيانات عن ست سنوات فقط. ولتوضيح الحسابات، نحصل على القيمة لـ$Z_{3(1969)}$ عن طريق حساب:

$$\begin{aligned} Z_{3(1969)} &= Y_{d(1968)} + 4Y_{d(1967)} + 9Y_{d(1966)} + 16Y_{d(1965)} \\ &= 590 \quad + \quad 2{,}188 \quad + \quad 7{,}568 \quad = \quad 14{,}954. \end{aligned}$$

وباستخدام البيانات الجديدة في الجدول رقم (٥-٣)يمكننا استخدام طريقتنا العادية لتقدير المعادلة:

$$C_t = a + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + u_t. \tag{5.39}$$

وتكون المعادلة المقدرة هي:

$$\hat{C}_t = -43.5 + 1.02 Z_{1t} - 1.44 Z_{2t} + 0.35 Z_{3t}. \tag{5.40}$$

وبمساعدة قيمنا المقدرة لـα_i، يمكننا حساب تقديراتb_i:

$$\begin{aligned} \hat{b}_0 &= \hat{\alpha}_0 = 1.02, \\ \hat{b}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 = 1.02 - 1.44 + 0.35 = -0.07, \\ \hat{b}_2 &= \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2 = 1.02 + 2(-1.44) + 4(0.35) = -0.46, \\ \hat{b}_3 &= \hat{\alpha}_0 + 3\hat{\alpha}_1 + 9\hat{\alpha}_2 = 1.02 + 3(-1.44) + 9(0.35) = -0.15, \\ \hat{b}_4 &= \alpha_0 + 4\hat{\alpha}_1 + 16\alpha_2 = 1.02 + 4(-1.44) + 16(0.35) = 0.86. \end{aligned}$$

وهكذا تكون معادلتنا المقدرة للاستهلاك ذات فترات إبطاء آلمون الأربع هي: *

* في هذه الحال، لاتتوافق المعادلة المقدرة توافقا جيدة مع توقعاتنا. وتوحي العلامات الجبرية لمعاملات القيم المبطأة لمتغير الدخل ضرورة التفكير جيداً عند صياغة دالة الاستهلاك.

$$\hat{C}_t = -43.5 + \underset{(3.3)}{}\ \underset{(5.3)}{1.02Y_{dt}} - \underset{(0.9)}{0.07Y_{d(t-1)}} + \underset{(2.8)}{0.46Y_{d(t-2)}}$$

$$- \underset{(1.9)}{0.15Y_{d(t-3)}} + \underset{(4.3)}{0.86Y_{d(t-4)}} \qquad R^2 = 0.99 \qquad (5.41)$$

جدول رقم (٥-٣)

Year	Ct	Z_{1t}	Z_{2t}	Z_{3t}
1964	401	1,942	3,667	10,821
1965	433	3,065	3,859	11,347
1966	466	2,213	4,104	12,030
1967	492	2,375	4,392	12,826
1968	537	2,560	4,742	13,860
1969	576	2,752	5,112	14,954

إضافة إلى القيم المقدرة للمعاملات، فقد عرضنا، أيضا، القيم المطلقة لنسب t، ولمعامل التحديد. وتظهر معظم برامج الحاسوب التي تتضمن استخدام منهج آلمون للتقدير هذه المعلومات الإضافية.

(٥-٢) استخدام المتغيرات الصورية

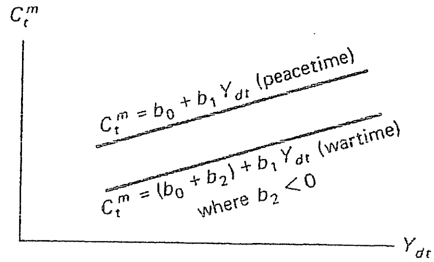
حتى الآن، تعاملنا تعاملا مكثفا مع المتغيرات التي يمكن قياسها في وحدات كمية، مثل مستوى الدخل المتاح، أو معدل التغير في معدلات الأجور. ولكن في كثير من الأحيان، نعتقد بوجود متغيرات معينة ذات أهمية عظمى ولكن، من طبيعة نوعية. فقد نعتقد، على سبيل المثال، أن مستوى الإنفاق الاستهلاكي الإجمالي يعتمد ليس فقط على مستوى الدخل المتاح، ولكن، أيضا، على ما إذا كان المجتمع في حالة حرب أم في حالة سلم. ولكن كيف ندخل متغيرا، لوقت السلم أو وقت الحرب في معادلة الانحدار.*

* لمناقشة أكثر تعمقاً من المناقشة التالية، انظر:

Arther Goldberger. *Econometric Theory* (New York: Wiley, 1964), pp. 218-227.

إحدى طرق مواجهة هذه المشكلة هي تقدير معادلتين منفصلتين للاستهلاك، وذلك عن طريق استخدام بيانات فترة الحرب لتقدير دالة الاستهلاك في وقت الحرب، واستخدام بيانات وقت السلم لتقدير دالة الاستهلاك في وقت السلم، وهكذا فسنحصل على معادلتين مختلفتين للاستهلاك. ولكن، هناك طريقة أكثر كفاءة وهي تقدير معادلة واحدة للفترتين ولكن بعد وضع بعض الافتراضات.

دعنا نفترض أن القيود على الاستهلاك وقت الحرب لاتغير الميل الحدي للاستهلاك، ولكنها تخفض الميل المتوسط له. ووفقا للشكل (٥-٥) نفترض أن دالة الاستهلاك خلال سنوات الحرب لها الميل نفسه كما في سنوات السلم، ولكن لها قاطع أدنى (أو حد ثابت أصغر). باستخدام هذ الافتراض نستطيع التعبير عن دوال الاستهلاك وقت الحرب ووقت السلم بدلالة معادلة انحدار واحدة هي:



الشكل رقم (٥-٥)

$$C_t = b_0 + b_1 Y_{dt} + b_2 D_t + u_t, \qquad t = 1, \cdots, n, \qquad (5.42)$$

حيث:

$D_t = 0$ لسنوات السلم

$D_t = 1$ لسنوات الحرب

تبين المعادلة (5.42) أنه خلال سنوات السلم عندما تكون $D_t = 0$ يصبح لدينا:

$$C_t = b_0 + b_1 Y_{dt} + u_t, \qquad (5.43)$$

بينما في فترات الحرب، عندما تكون $D_t = 1$ فإنه يصبح لدينا:

$$C_t = (b_0 + b_2) + b_1 Y_{dt} + u_t, \tag{5.44}$$

حيث يفترض أن $b_2 < 0$.

افترض أن الفترة الزمنية موضع الاهتمام هي حيث تكون فترة سنوات الحرب من $t=5$ إلى $t=9$. في هذه الحال، وفقا للمعادلة (5.42)، لدينا مجموعة من البيانات مثل الموجودة في الجدول رقم (٥-٤) وباستخدام هذه البيانات يمكننا تقدير قيم المعاملات في المعادلة (5.42) بطريقة الانحدار المتعدد العادية.

افترض أننا فعلنا ذلك وحصلنا على المعادلة:

$$\hat{C}_t = 40 + 0.9 Y_{dt} - 30 D_t, \tag{5.45}$$

جدول رقم (٥-٤)

t	الدخل المتاح	الإنفاق الاستهلاكي	D
1	Y_{d1}	C_1	0
.	.	.	.
.	.	.	.
.	.	.	.
4	Y_{d4}	C_4	0
5	Y_{d5}	C_5	1
5	Y_{d6}	C_6	1
7	Y_{d7}	C_7	1
8	Y_{d8}	C_8	1
9	Y_{d9}	C_9	1
10	Y_{d10}	C_{10}	1
.	.	.	.
.	.	.	.
.	.	.	.
n	Y_{dn}	C_n	0

حيث يتضح من نسبة t المناظرة إلى المتغير D_t أنها ذات حجم كاف مما يوحي بأن المعلمة b_2 في المعادلة (5.42) غير صفرية. ومن ثم يجب أن نستنتج بأن الحرب لها تأثير سالب ومعنوي على الإنفاق الاستهلاكي. وستصبح معادلة الاستهلاك المقدرة لسنوات السلم وسنوات الحرب هي على الترتيب.

$$\hat{C}_t = 40 + 0.9Y_{dt} \tag{5.46}$$

و

$$\hat{C}_t = 10 + 0.9Y_{dt}, \tag{5.47}$$

فإذا قيست النفقات الاستهلاكية ببلايين الدولارات، فإن المقارنة بين المعادلة (5.46) و المعادلة (5.47) توضح أنه، عند مستويات الدخل المختلفة، تنخفض النفقات الاستهلاكية بمقدار 30 بليون دولار خلال سنوات الحرب. ويوضح الشكل رقم (٥-٦) هـذه الدوال، حيث نرى أن دالة الاستهلاك لوقت الحرب DE هي خط مستقيم بميل دالة الاستهلاك نفسه لوقت السلم AB ولكن مع قاطع رأسي يقل بمقدار ٣٠ عن AB.

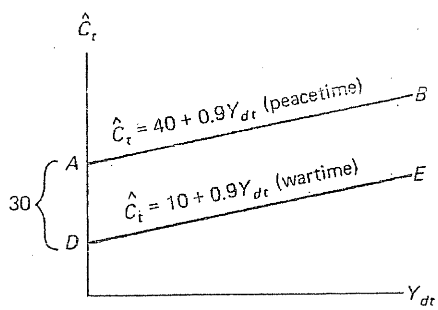
وبالمقابل، قد نفترض أن ظروف وقت الحرب تقلل الميل الحدي للاستهـلاك دون الحد الثابت في معادلة الاستهـلاك.* في هذه الحال، تأخذ معادلتنا للانحدار المشتملة على كلتا الفترتين الشكل:

$$C_t = b_0 + b_1Y_{dt} + b_2(Y_{dt}D_t) + u_t, \tag{5.48}$$

حيث لدينا مرة أخرى

$D_t = 0$ لسنوات السلم،

$D_t = 1$ لسنوات الحرب.



شكل (٥-٦)

* في التطبيق، يمكن للفرد أن يقرر ما إذا كان القاطع أو بالمقابل، الميل الحدي للاستهلاك هو الذي ينتقل عن طريق دراسة أشكال أنواع القيود المفروضة على الاستهلاك وقت الحرب ... إلخ.

تبين المعادلة (5.48) بأن دالة الاستهلاك في أوقات السلم هي:

$$C_t = b_0 + b_1 Y_{dt} + u_t, \tag{5.49}$$

وذلك طالما أن $D_t = 0$، بينما تكون في أوقات الحرب

$$C_t = b_0 + (b_1 + b_2) Y_{dt} + u_t, \tag{5.50}$$

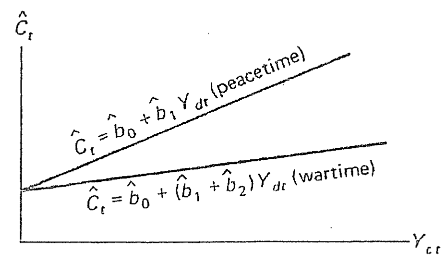
ونتوقع أن $b_2 < 0$. وكما سبق، فإنه يمكن استخدام بيانات مثل تلك الموجودة في الجدول رقم (٥-٤) لتقدير المعادلة (5.48). وستشابه العلاقة المقدرة الناتجة تلك المنحنيات الموجودة في الشكل رقم(٥-٧)، حيث يكون لدالة الاستهلاك في وقت الحرب انحدار أقل، ولكن القاطع الرأسي نفسه كما هو الحال في وقت السلم.

ويطلق على المتغير D الذي يظهر في المعادلة أعلاه «المتغير الصوري» dummy variable، وهو يأخذ قيمة الواحد الصحيح إذا تحققت بعض الشروط وقيمة الصفر إذا لم تحقق. وفي الحقيقة فإن استخدام المتغيرات الصورية تعميم قوي جدا لتحليل الانحدار، وكما سنري فإنه يسمح لنا تعميم مجال تحليلنا للإنحدار ليشمل متغيرات مهمة لايمكننا وصفها في وحدات كمية. وهكذا فباستخدام المتغيرات الصورية يمكن الأخذ في الحسبان تأثير العوامل النوعية المهمة التي تؤثر على متغيرنا التابع.

اعتبر، مرة أخرى، المثال الأول أعلاه، حيث افترضنا تغير القاطع وثبات الميل الحدي للاستهلاك (م ح س) خلال سنوات الحرب، وبدلا من استخدام طريقة المتغيرات الصورية، قمنا بعمل دالتين للاستهلاك، واحدة لسنوات الحرب وأخرى لسنوات السلم. في هذه الحال، سنقدر أربع معلمات وليس ثلاثا. وفي ظل افتراضنا بأن (م ح س) هو نفسه لوقت الحرب ووقت السلم فسوف ننتهي بتقديرين لمعلمة واحدة. وتصبح مشكلتنا هي استخدام هذين التقديرين للحصول على تقدير وحيد «أفضل» لـ (م ح س). في مثل هذه الحال، سنحاول، على الأرجح، تطوير بعض الطرق للحصول على القيمة المتوسطة لهذين التقديرين.

يمكن إثبات أنه إذا دمج هذان المتغيران معا بطريقة مثلى (يجب تحديدها) فإن النتيجة تكون متماثلة مع التقدير الذي يتحقق بوساطة طريقة المتغير الصوري تحققا أكثر دقة، دع b_{1p} هو مقدر الـ (م ح س) المبني على معادلة وقت السلم وبياناته، دع،

b_{1w} أيضا، المقدر المناظر المشتق من معادلة وقت الحرب وبياناته. في ظل تحقيق افتراضاتنا العادية، فإن هذه المقدرات غير متحيزة. فإذا دمجت هذه المقدرات دمجا ينجم عنه مقدر غير متحيز لـ (م ح س) مع أصغر التباينات الممكنة، فإن المقدر الناتج يكون متماثلا مع المقدر الذي يتحقق بوساطة المتغير الصوري. وفي الحقيقة، فإن طريقة المتغير الصوري تستخدم معلومات العينة المتاحة كافة - إضافة إلى المعلومات المسبقة المرتبطة بتغيرات المعلمة - بأفضل الطرق الممكنة.



شكل (٥-٧)

ويمكننا - بهذه الناسبة - استخدام عدد كبير من المتغيرات الصورية بقدر مانريد، شرط أن يتوافر لدينا عدد كاف من المشاهدات يسمح بتقدير المعادلة. افترض، على سبيل المثال، أننا نرغب في تفسير السلوك الاستهلاكي لأسر مختلفة. وأننا نعتقد أن مستوى استهلاك هذه الأسر يعتمد على عدد من السمات المميزة لها، كوجود الأطفال أو غيابهم، وما إذا كانت الأسرة تقيم في منزل تملكه أو تستأجره. وعنصر أرباب الأسر، وهلم جرا، إضافة إلى الدخل المتاح. فإذا استطعنا الحصول على هذه المعلومات كافة لعينة من الأسر، فإنه يمكن، على سبيل المثال، تقدير المعادلة:

$$C_t = b_0 + b_1 Y_{dt} + b_2 F + b_3 H_t + b_4 R_t + b_5 A_t + u_t, \qquad (5.51)$$

حيث:

C_t = الإنفاق الاستهلاكي للأسرة t.

Y_{dt} = الدخل المتاح للأسرة t.

F_t = 1 إذا كانت الأسرة لها أطفال.

0 إذا كانت الأسرة بدون أطفال.

H_t = 1 إذا كانت الأسرة تملك المنزل الذي تقيم فيه.

0 إذا كانت الأسرة لاتملكه

R_t = 1 إذا كانت الأسرة من العنصر الأبيض

0 إذا كانت الأسرة غير ذلك

A_t = 1 إذا كانت رب الأسرة يتجاوز الخمسين عاما

0 إذا كان رب الأسرة غير ذلك.

u_t = الخطأ العشوائي.

مثال

إن تطبيقات المتغيرات الصورية، في الحقيقة غير محدودة. دعنا نأخذ مثالا آخر يتضمن نوعا مختلفا تماما من المشاكل، والمثال من دراسة حديثة لأحد مؤلفي هذا الكتاب.* والقضية موضوع الدراسة هي ما إذا كانت الطبيعة الرسمية للدستور السياسي للدولة لها تأثير منتظم على درجة اللامركزية في المالية العامة للدولة. أو باختصار، هل الدستور عامل مهم في تحديد النصيب النسبي للنشاط المالي للحكومة المركزية في القطاع العام ككل؟

بعد أن تحققنا من أهمية متغيرات أخرى كحجم السكان، ومستوى الدخل لكل نسمة، فإن الطريقة هي إدخال متغير صوري بقيمة الواحد الصحيح إذا كانت

* أنظر: W. E. Oates. *Fiscal Federalism* (New York: Harcourt Brace Jovanovich, 1972), Chap. 5.

للدولة دستور فيدرالي (أي يضمن بعض الاستقلال الذاتي للمستويات الحكومية المحلية) أو قيمة الصفر في غياب الدستور الفيدرالي (حيث يحدد مجال السلطات الحكومية المحلية بوساطة الحكومة المركزية). باستخدام البيانات المقطعية لعينة من ٥٣ دولة، كانت المعادلة المقدرة هي:

$$\begin{aligned} \hat{G} = \ & 96 \quad - 1.21 \ln P - 0.004Y - 0.6Z - 15.9F \qquad N = 53, \\ & (12.1) \quad (1.3) \quad (2.3) \quad (5.5) \quad (4.7) \qquad R^2 = 0.65 \end{aligned} \tag{5.52}$$

(حيث إن الأرقام الموجودة داخل الأقواس تحت المعاملات المقدرة هي القيم المطلقة لنسبة t)، وإن:

G = نصيب الحكومة المركزية من الإيرادات العامة الكلية (بالنسبة المئوية)،

ln P = اللوغاريتم الطبيعي لعدد السكان (بالآلاف)،

Y = متوسط دخل الفرد بالدولارات الأمريكية ١٩٦٥م.

Z = مساهمات الضمان الاجتماعي بوصفها نسبة مئوية من إجمالي الإيرادات العامة الجارية،

F = 1 للدول ذات الدساتير الفيدالية،

0 للدول ذات الدساتير غير الفيدرالية.

تتفق نتائج المعادلة (5.52) بوضوح مع الافتراض بأن وجود دستور فيدرالي يسهم في زيادة درجة اللامركزية في الماليات العامة. ويظهر معامل المتغير الصوري F بإشارة سالبة، كما أن نسبة t المناظرة له تزيد على أربعة. لذا، يمكننا أن نرفض، بسهولة، فرضية العدم القائلة بعدم وجود ارتباط بين G و F عند مستوى المعنوية 5%. كما يدل حجم المعامل بعد الأخذ في الحسبان تأثير حجم السكان، الدخل، وغيرها، على أن الحكومة المركزية في الدول الفيدرالية تحصل في المتوسط على 16% أقل من إجمالي الإيرادات العامة التي تحصل عليها الحكومات المركزية في الدول غير الفيدرالية بنسبة 16%. وهكذا تؤدي الفيدرالية المفروضة بوساطة الدساتير السياسية دورا مهما في تحديد درجة اللامركزية في النشاط التمويلي العام.

بعض النتائج الإضافية

بينت الدراسات أن المتغيرات الصورية مفيدة جدا في فصل الاختلافات الفصلية والإقليمية في السلوك. فمبيعات السيارات نتيجة لإدخال نماذج جديدة في فصل الأعياد، أو حجم الإنتاج من المحاصيل المختلفة التي تعتمد على ظروف الطقس، من الواضح أنها تتغير تغيرا منتظما مع فصول السنة، فإذا تعاملنا مع بيانات ربع سنوية أو شهرية فإنه يمكننا ادخال متغيرات صورية تناظر مختلف الفصول لنأخذ في الحسبان تأثيراتها. وبالمثل، عندما تتوقع وجود اختلافات إقليمية في السلوك، فإنه يمكننا أن نسمح بذلك عن طريق إدخال المتغيرات الصورية لمختلف الأقاليم.

ولكي نعرف كيف يتم ذلك (ولنشير إلى أحد المحاذير التي ينبغي تجنبها) اعتبر المثال التالي: افترض أننا نحاول فهم العلاقة بين حجم الإنتاج من إحدى السلع وحجم العمل المطلوب لإنتاجها حيث نعتقد بوجود قوى فصلية منتظمة تؤثر في هذه العلاقة. لذا، قد يمكننا وضع النموذج على النحو:

$$Q_t = b_0 + b_1 L_t + b_2 S_t + b_3 H_t + b_4 F_t + b_5 W_t + u_t, \qquad (5.53)$$

حيث:

Q_t = وحدات الإنتاج في الفصل (ربع السنة) t.

L_t = وحدات عنصر العمل

S_t = 1 للفصل السنوي ابريل - يونيو.
0 للفصول الأخرى

H_t = 1 للفصل السنوي يوليو - سبتمبر
0 للفصول الأخرى

F_t = 1 للفصل السنوي اكتوبر - ديسمبر
0 للفصول الأخرى.

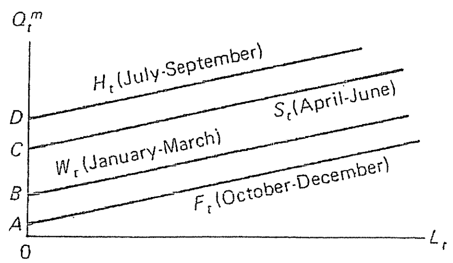
W_t = 1 للفصل السنوي يناير - مارس
0 للفصول الأخرى.

لاحظ أننا أدخلنا متغيرًا عشوائيًا لكل واحد من الفصول ربع السنوية الأربعة

في السنة الميلادية. وبدلالة الشكل رقم(٥-٨) يبين نموذجنا أن القيمة المتوسطة لمستوى الإنتاج، أو Q_t^m، تعادل $(b_0 + b_1 L_t)$ حيث إن b_1 هو الميل المشترك للخطوط الأربعة مضافا إليه كمية إضافية (يمكن أن تكون سالبة) تعتمد على الفصل ربع السنوي.

افترض أننا حاولنا تقدير المعادلة (5.53)، سوف نجد أنها، في شكلها العام، لايمكن تقديرها بسبب أنها تتضمن ارتباطا متعددًا خطيًا تامًا بين المتغيرات المستقلة. أي أن افتراضنا بأن المتغيرات المستقلة ليست مرتبطة فيما بينها تماما وخطية قد خولف وبالتحديد فإن لدينا:

$$S_t + H_t + A_t + W_t \equiv 1 \qquad (5.54)$$



شكل (٥-٨)

نعلم أنه أن أحد هذه المتغيرات، في أي فصل ربع سنوي معين، يأخذ قيمة الواحد الصحيح بينما البقية تأخذ قيمة الصفر، فمجموعهما ينبغي أن يكون الواحد دائما. تذكر من الفصل السابق أنه في وجود الارتباط الخطي المتعدد التام فلن نستطيع إيجاد تقديرات وحيدة للمعلمات لأنه لايتوافر لدينا عدد كاف من المعادلات الطبيعية.

ولكن هذه المشكلة يمكن حلها بسهولة عن طريق إسقاط أحد المتغيرات الصورية من المعادلة، وتغيير بعض تفسيراتنا. افترض، على سبيل المثال، أننا نسقط W_t من المعادلة (5.53) للحصول على:

$$Q_t = b_0 + b_1 L_t + b_2 S_t + b_3 H_t + b_4 F_t + u_t \,, \qquad (5.55)$$

والآن، فقد أزلنا الاعتماد الخطي بين المتغيرات المستقلة، وسنستطيع تقدير المعلمات

الخمس في المعادلة (5.55) خلال شهور الشتاء (يناير-مارس)، وستكون S_t، H_t و F_t مساوية الصفر، كما سيكون الحد الثابت لمعادلتنا هو b_0 أي أن b_0 ستناظر القاطع الرأسي OB في الشكل رقم(٥-٨). وعلى نحو مشابه، وبالرجوع إلى المعادلة (5.55)، نجد أن المعامل لكل متغير صوري يشير إلى كيفية اختلاف تأثير الفصل المناظر عن تأثر فصل الشتاء. على سبيل المثال، فخلال فصل الربيع (ابريل - يونيو)، عندما يأخذ المتغير S_t قيمة الواحد الصحيح، فإنه يكون لدينا:

$$Q_t = (b_0 + b_2) + b_1 L_t + u_t. \qquad (5.56)$$

وفي الشكل رقم(٥-٨) يساوي القاطع الرأسي للمنحنى والمناظر للفصل (أبريل-يونيو) -OC- $(b_0 + b_2)$ وعليه فإن b_2 يشير لاتجاه اختلاف أثر فصل الربيع عن فصل الشتاء وحجمه. فمثلا في الشكل رقم(٥-٨)يتوقع أن تكون b_2 موجبة وبالمثل من الشكل ذاته يتوقع أن تكون $0 > b_3$ و $0 > b_4$.

والآن، نحن في وضع يسمح لنا بمراجعة الأشياء. فإذا كان لدينا أربعة فصول واعتقدنا أن مستوى معادلتنا يتغير وفقا لكل فصل، نأخذ أحد هذه الفصول ونعده فصلنا المعياري. وننسب تأثير الفصول الأخرى لتأثيره. ففي المثال السابق اخترنا فصل الشتاء ليكون الفصل المعياري، أما إذا أسقطنا S_t وضمنا H_t، F_t و W_t في معادلتنا، يصبح فصل الربيع، في هذه الحال، هو الفصل المعياري. وهنا، ستمثل b_0 ارتفاع المعادلة خلال فصل الربيع. وبمعنى آخر، يمكن القول إنه إذا كان لدينا أربعة فصول واعتقدنا بأن القاطع الرأسي لمعادلتنا يتغير وفقا لكل فصل، فإننا نكون معادلة تحتوي على أربع معلمات تصف هذه القواطع للمعادلة. ولما كان القاطع هو أحد هذه المعلمات فإننا نحتاج فقط إلى ثلاثة متغيرات صورية. لاحظ أننا لم نستطع تقدير المعلمات لمعادلتنا الأصلية (5.53) لأن هذه المعادلة تحتوي على خمس معلمات للقاطع هي: b_5, b_4, b_3, b_2, b_0. ولكن، هناك فقط أربعة قواطع متعلقة بالفصول الأربعة. ولذا فإن احد المتغيرات الصورية لاداعي له. ويصبح تعميمنا: إذا توقعنا أن التغيرات الإقليمية أو الفصلية تسبب k مستويات مختلفة من المعادلة فإننا نحتاج لاستخدام (k-1) فقط من المتغيرات الصورية. وحتى تستوعب هذه المناقشة على نحو

أفضل، يمكنك العودة إلى معادلة الاستهلاك التي استخدمناها في بداية هذا المبحث، واعتبر مايمكن أن يحدث فيما لو وضعنا في المعادلة (5.42) متغيرًا صوريًا لسنوات الحرب:

W = 1 خلال سنوات الحرب

0 خلال السلم،

ومتغيرًا صوريًا ثانيا لسنوات السلم

P = 1 خلال سنوات السلم،

0 خلال أوقات الحرب.

وبالتحديد، يجب أن تكون قادرا على إثبات أنه إذا أدخلنا كلا من W و P في المعادلة، فإنه لايمكن تقديرها. كما ينبغي أن تبرهن بأن معادلة واحدة تغطي كلا من فترات الحرب والسلم يمكن تكوينها بأي من W أو P.

(٥-٣) الشكل الدالي مرة أخرى

في الفصل الثالث، فحصنا عددا من التحويلات التي تمكننا من وضع العلاقات غير الخطية في شكل خطي، وذلك حتى يمكننا استخدام نموذج الانحدار الخطي العادي. يمكن تعميم استخدام هذه التحويلات بسهولة، ففي حالة الانحدار المتعدد، لن نتعمق في هذا الموضوع، وبدلا من ذلك، سنأخذ مثالا واحدا معينا - التحويل اللوغارتمي - لنرى كيف نوسع تحليلنا من حالة نموذج انحدار المتغيرين إلى حالة الانحدار المتعدد. وحينئذ، سنطور تحويلات إضافية أخرى مفيدة تصبح ممكنة عندما لانكون مقيدين بحالة انحدار المتغيرين السابق.

التحويل اللوغارتمي المعمم

تذكر من الفصل الثالث أننا اختبرنا علاقة إنتاجية بسيطة من الشكل:

$$Q_t = aL_t^b e^{u_t} , \qquad (5.57)$$

حيث L_t هي كمية العمل المستخدمة في الفتـرة t، وهو عنصر الإنتاج الوحيد فـي إنتاج Q_t. وبأخذ اللوغاريتمات في المعادلة (5.57)، نعبر عن دالة الإنتاج هذه على النحو:

$$\ln Q_t = \ln a + b \ln L_t + u_t . \tag{5.58}$$

بعدئذ، نحول المعادلة (5.58) إلى الشكل رقماالخطي:

$$Q_t^* = a^* + bL_t^* + u_t , \tag{5.59}$$

عن طريق وضع:

$$Q_t^* = \ln Q_t ,$$
$$a* = \ln a ,$$
$$L_t^* = \ln L_t .$$

في هذا الشكل، رقماستخدمنا منهجنا العادي للتقدير للحصول على المقدرات غير المتحيزة $\hat{a}^*$، $\hat{b}$ للمعلمات في المعادلة (5.59)، وبعدئذ، أخذنا $e^{\hat{a}^*}$ مقدراً متحيزاً ولكن متسقًا لـ a.

أحد القيود الواضحة في المعادلـة (5.57) هو اشتمالها على عامل متغيـر واحد للإنتاج، ذلك أن السلع والخدمات تنتج، عادة، باستخدام تشكيلـة مـن المدخلات. ويوحي هذا بأن دالة الإنتاج الأكثر واقعية ينبغي أن تشتمل عـلـى كميات متغيرة من عوامل إنتاجية متعددة. لتحقيق ذلك دعنا نأخـذ شكلا عاما للمعادلة (5.57):

$$Q_t = b_0 F_{1t}^{\ b_1} F_{2t}^{\ b_2} \cdots F_{kt}^{\ b_k} e^{u_t} , \tag{5.60}$$

حيث يمثل كل من F_t's كمية من عامل إنتاجي معين يستخدم خلال الفتـرة t. على سبيل المثال، قد تشير F_{1t} إلى كمية العمل المستخدمة في الفترة t، بينما تشير F_{2t} إلى

* تحتوي دالة الإنتاج لكوب ودوجلاس على عدد من السمات الملائمة والمهمة جعلتها ذات استخدام كبير في التحليل الاقتصادي. لمناقشة هذه السمات، يرجع إلى:

James M. Henderson and Richard E. Quandt, *Micro-economic Theory*, 2nd ed. (New York: McGraw-Hill, 1971), pp. 79-85.

مساحة الأرض وهلم جرا. وعلاقة الإنتاج التي تأخذ مثل هذا الشكل رقمتعرف بدالة إنتاج كوب - دوجلاس Cobb-Douglas. * ومانحتاج معرفته هو قيم b's في هذه العلاقة حتى يمكننا أن نعرف كيف يتغير الناتج الكلي مع تغير كميات المدخلات المختلفة. فقد نكون مهتمين، على سبيل المثال، بمعرفة ما إذا كان إنتاج سلعة معينة يخضع لظاهرة تزايد غلة الحجم. ومعنى ذلك أنه مع ثبات الأشياء الأخرى*، إذا استطعنا مضاعفة المدخلات من جميع عناصر الإنتاج الأخرى، هل يزداد الإنتاج بأكثر من الضعف؟ فإذا كان الأمر كذلك فإننا نصرح بأن لدينا تزايدا في غلة الحجم، أما إذا تزايد الإنتاج إلى الضعف بالضبط، حينئذ هناك ثبات في غلة الحجم، وأخيرا إذا تزايد الإنتاج بأقل من الضعف، يوجد لدينا حينئذ تناقصا في غلة الحجم.

ويمكن تحديد ذلك بسهولة من دالة إنتاج كوب-دوجلاس عن طريق أخذ المجموع $\Sigma_{k=1}^{i} b_i$. ويتضح لنا ذلك بأخذ مثال مبسط، ونترك التعميم للقارئ على سبيل التمرين. افترض أن لدينا سلعة، Q، تنتج باستخدام عنصري العمل ورأس المال فقط على النحو:

$$Q_t = b_0 L_t^{b_1} K_t^{b_2} e^{u_t}, \tag{5.61}$$

حيث L_t وK_t هما كميات العمل ورأس المال، على التوالي والمستخدمة في إنتاج Q خلال الفترة t. والآن افترض أنا ضاعفنا مدخلات العمل ورأس المال، دع Q'_t هو المستوى الجديد من الإنتاج، حينئذ، سيكون لدينا:

$$\begin{aligned} Q'_t &= b_0 (2L_t)^{b_1} (2K_t)^{b_2} e^{u_t} = b_0 (2^{b_1})(L_t^{b_1})(2^{b_2})(K_t^{b_2}) e^{u_t} \\ &= 2^{(b_1+b_2)} b_0 L^{b_1} K_t^{b_2} e^{u_t} = 2^{(b_1+b_2)} Q_t. \end{aligned} \tag{5.62}$$

يتضح لنا من المعادلة الأخيرة في (5.62) أنه إذا كانت لدينا $1 < (b_1 + b_2)$ فإن الإنتاج سوف يزداد بأكثر من الضعف وسيكون لدينا تزايد في غلة الحجم، وإذا كانت $b_1 + b_2 = 1$ يتضاعف الإنتاج بالضبط، فهناك ثبات في غلة الحجم، وأخيرا

* يرتبط الشرط «مع بقاء الأشياء الأخرى على حالها» بالخطأ العشوائي، أي أننا نفترض، في المناقشة أن الخطأ العشوائي لايتغير عندما تتغير قيم المدخلات.

إذا كانت $b_1 + b_2 < 1$ فإن الإنتاج يزداد بأقل من الضعف حيث يوجد تناقص في غلة الحجم. وبعمومية أكثر في المعادلة (5.60) ينبغي أن تكون قادراً على إثبات أن حالات تزايد غلة الحجم وثباتها وتناقصها تناظر الحالات التي يزيد فيها $\sum_{i=1}^{i} b_i$ عن الواحد، يساوي الواحد أو يقل عن الواحد الصحيح على التوالي*. وعمليا تصبح مشكلتنا هي تقدير قيمة الـ b's من أجل تحديد طبيعة العلاقة الإنتاجية لسلعة معينة. وللحصول على المعادلة (5.60)، في شكل قابل للتقدير نستخدم التحويل اللوغارتمي، فبأخذ لوغارتمات المعادلة (5.60)، نحصل على:

$$\ln Q_t = \ln b_0 + b_1 \ln F_{1t} + b_2 \ln F_{2t} + \cdots + b_k \ln F_{kt} + u_t . \qquad (5.63)$$

بعدئذ، نعرف:

$$\begin{aligned} Q_t^* &= \ln Q_t , \\ b_0^* &= \ln b_0 , \\ F_{it}^* &= \ln F_{it} . \end{aligned}$$

وبالتعويض في المعادلة (3.63)، يصبح لدينا مايلي:

$$Q_t^* = b_0^* + b_1 F_{1t}^* + b_2 F_{2t}^* + \cdots + b_k F_{kt}^* + u_t . \qquad (5.64)$$

وباستخدام طريقتنا العادية في التقدير على المتغيرات المعرفة في المعادلة (5.64)، يمكننا الحصول على مقدرات غير متحيزة $\hat{b}_0$، $\hat{b}_1$، ... و $\hat{b}_k$ لمعلمات المعادلة (5.64)، وبهذه الطريقة نحصل على مقدرات غير متحيزة لمرونات الناتج لكل عنصر من عناصر الإنتاج. وكما هو في حالة الانحدار البسيط، سيكون $\hat{b}_0 = e^{\hat{b}_0^*}$. وأخيرا، وباستخدام النتائج الموضحة في الملحق B لهذا الفصل، يمكننا اختبار فرضية وجود ثبات غلة الحجم.

---

* إحدى السمات الأخرى المفيدة لدالة الإنتاج كوب - دوجلاس في المعادلة (5.60) هي أن كل b_i يمكن تفسيرها على أنها مرونة الناتج بالنسبة للعامل i. أي إذا كانت F_i تتزايد بنسبة ١٪، وجميع المدخلات الأخرى تظل ثابتة، فإن الناتج Q سيزداد بنسبة b_i في المائة. لكن، على القارئ أن يلاحظ أن كل متغير لايمكن الاستغناء عنه في عملية الإنتاج، بمعنى أنه إذا كانت $F_i=0$، فإن الناتج Q سيصبح صفرا أيضا.

أشكال متعددات الحدود للمتغيرات المستقلة

يود الاقتصاديون، عادة، التعامل مع إمكانية أن تكون العلاقات بين المتغيرات الاقتصادية غير خطية إلا أنهم غير متأكدين، في الحقيقة، من شكل هذه العلاقة. فمثلا، اعتبر تأثير عمر الفرد على إنفاقه الاستهلاكي. فمن الممكن أنه، مع تقدم عمر الفرد واتساع خبراته، أن تحفزه معلوماته عن الأنشطة المختلفة على زيادة إنفاقه على السلع والخدمات الاستهلاكية. ولكن، بعد وصوله إلى عمر معين، قد يتباطأ إنفاق الفرد، وفي الحقيقة، يبدأ الفرد في تخفيض مستوى إنفاقه الاستهلاكي. مثل هذه العلاقة بين الإنفاق الاستهلاكي والعمر (عندما تظل المتغيرات الأخرى الملائمة كمستوى الدخل ثابتة) تظهر في المنحنى المتصل A في الشكل رقم(٥-٩).

من ناحية، أخرى قد نجد أنه، مع تقدم الفرد في العمر، يزداد طلبه على الأمن الاقتصادي ومن ثم، على الادخار، ولذا، يتناقص إنفاقه الاستهلاكي بانتظام. مثل هذه العلاقة تظهر في الخط المتقطع B في الشكل نفسه. وأخيرا قد يتناقص إنفاق الفرد الاستهلاكي في البداية ببطء شديد وبعدئذ، كلما تقدم الفرد في العمر يتناقص استهلاكه بمعدل متزايد. وتظهر هذه العلاقة في الخط المنقط C في الشكل رقم(٥-٩).

وعلى سبيل أحد الأمثلة الأخرى للمتغيرات التي ترتبط بعضها بعضا بعلاقة أقرب إلى عدم الخطية، اعتبر تأثير التغيرات في تكلفة المعيشة على تعديلات الأجور، كما تقاس بالتغيرات في الأجور النقدية. وبالطبع يمكن جعل التغيرات في تكلفة المعيشة تنعكس تماما في تغيرات الأجور. أي أنه، مع بقاء الأشياء الأخرى، على حالها، إذا ارتفعت تكلفة المعيشة بمقدار X في المائة ارتفع معدل الأجور بمقدار X في المائة*. ومن الناحية الأخرى من المحتمل أن التغيرات الطفيفة في تكلفة المعيشة لاتلاحظ ومن ثم لاتؤدي إلى زيادات مناظرة في الأجور. في هذه الحال، قد نفترض أن التغيرات الكبيرة في تكلفة المعيشة هي فقط التي تنعكس في تعديلات الأجور. تظهر هذه الاحتمالات في منحنيات OA_1 و OA_2A_3 على الترتيب في الشكل رقم(٥-١٠).

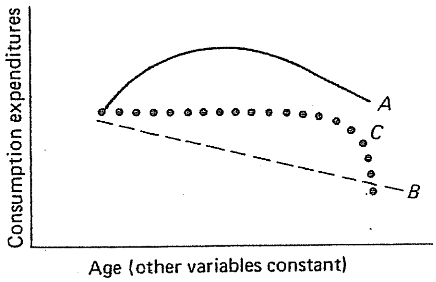
---

* قد ترغب في إضافة بعض الزيادات الإضافية في الأجور لتعكس الزيادة في الإنتاجية.

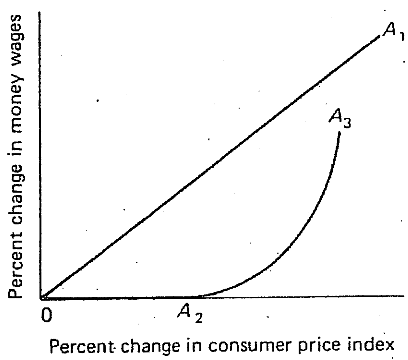
وفي ضوء هذه الأمثلة، نرغب في توجيه اهتمامنا نحو مشاكل التقدير واختبار العلاقة بين المتغيرات عندما لايكون شكل هذه العلاقة مؤكدا. سنقيم نتائجنا من خلال عرض أمثلة إضافية في مبحث آخر من هذا الفصل.

نبدأ (استخدام حالة المتغيرين للتبسيط) بافتراض أن المتغير التابع Y يرتبط بالمتغير المستقل X ارتباطا غير مؤكد. هذا الافتراض يمكن التعبير عنه على النحو:

$$Y_t = f(X_t) + u_t \,, \tag{5.65}$$



شكل (٥-٩)



شكل (٥-١٠)

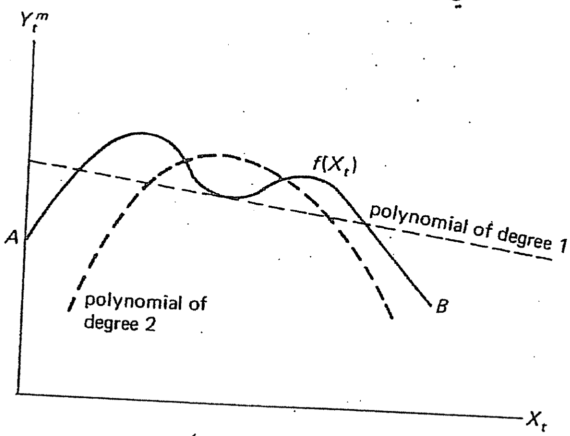
حيث إن u_t هو الخطأ العشوائي. تبين المعادلة (5.65)، ببساطة أن القيمة رقم t لـ Y أو Y_t، تعتمد على القيمة رقم t لـ X أو X_t، وعلى الخطأ العشوائي u_t. ولأننا لانعلم الشكل المحدد لـ $f(x_t)$ في المعادلة (5.65)، فإنه لتحقيق بعض التقدم في تقدير العلاقة بين Y_t و X_t ينبغي علينا إما أن نوجد شكل $f(x_t)$ أو بالمقابل، استخدام بعض التقريب لها. سوف نستخدم المنهج الأخير عن طريق استدعاء النظرية التي ذكرناها واستخدمناها عند الحديث عن طريقة فترة الإبطاء لآلمون. وبالتحديد، تصرح النظرية، أنه في ظل ظروف عامة، فإن دالة (أو منحنى) قد تقرّب لأي درجة من الدقة بوساطة متعدد للحدود. فإذا كان الأمر كذلك فيمكننا أن نطبق هذه النظرية على دالتنا غير المعلومة في المعادلة (5.65) للحصول على:

$$f(X_t)=a_0+a_1X_t+a_2X_t^2+\cdots+a_kX_t^k. \qquad (5.66)$$

وعمومًا، كلما أردنا درجة أكبر من الدقة ينبغي علينا استخدام درجة أعلى من متعدد الحدود (k). ينتج هذا من مناقشتنا بمتعدد الحدود في طريقة فترة الإبطاء لآلمون. في ذلك المبحث، اشرنا إلى أن عدد نقط الانقلاب على شكل متعدد الحدود تقل عن درجة متعدد الحدود بمقدار واحد على الأكثر. من هذا، قد نستخلص أن متعدد الحدود ذا الدرجات الأعلى يكون أكثر مرونة من ذلك المرتبط بدرجات أقل. ويشير هذا إلى أننا إذا أردنا تقريبا أدق فإننا نحتاج إلى متعدد للحدود ذي درجة أعلى حتى يتضمن مرونة كافية تتبع اتباعا وثيقا شكل الدالة غير المعلومة. وبالتدريج، فكلما كان شكل الدالة التي نرغب في تقريبها معقدا أزدادت بالضرورة درجة متعدد الحدود المناظر.

وقد وضّح ذلك بدلالة الدالة غير المعلومة $f(x_t)$ في الشكل رقم(٥-١١). في هذا الشكل نفترض أن دالتنا غير المعلومة لها الشكل المشار إليه بالخط المتصل AB. الآن يمكن تعريف هذا المنحنى - على الرغم من أن ذلك يتم بطريقة ضعيفة - بخط مستقيم يترتب على استخدام متعدد للحدود من الدرجة الأولى (k=1). ويتحسن التقريب إذا ما استخدمنا بدلا من ذلك متعددا للحدود من الدرجة الثانية (k=2). ويمكن أن يتحسن التعريف أكثر عن طريق k=3 وهلم جرا.

افترض أنه يتوافر لدينا عدد من الافتراضات المرتبطة بالشكل رقمالعام للعلاقة بين Y_t و X_t. افترض، أيضا، أن أكثر هذه الافتراضات تعقيدا تقترح أن قيمة $k = k_m$. ستكون ملائمة للتقريب. في المعادلة (5.66)*، ونعني بكلمة «ملائمة» أن علامة تقريبا مساوية لـ $\doteq$ في المعادلة (5.66) قد تستبدل بخسارة قليلة من الدقة بعلامة التساوي. نلاحظ، أيضا، أنه إذا كان متعدد الحدود من الدرجة k_m هو تقريب ملائم لأكثر أشكالنا المفترضة تعقيدا، فإنه يكون تقريبا ملائما، أيضا، للأشكال الأبسط كافة التي نعتبرها.



شكل (٥-١١)

في ظل هذا الافتراض المرتبط بـk_m، قد نستخدم $k = k_m$ في المعادلة (5.66) وهذا، بدوره، في المعادلة (5.65) للحصول على:

$$Y_t = a_0 + a_1 X_t + a_2 X_t^2 + \cdots + a_{k_m} X^{k_m} + u_t. \qquad (5.67)$$

ويمكن تحويل المعادلة (5.67) إلى نموذجنا العادي عن طريق التعويضات:

$$Z_{it} = X_t^i, \qquad i = 1, \cdots, k_m \qquad (5.68)$$

أي إذا قمنا بالتعويض من المعادلة (5.68) في المعادلة (5.67) سنحصل على:

$$Y_t = a_0 + a_1 Z_{1t} + a_2 Z_{2t} + \cdots + a_{k_m} Z_{k_m t} + u_t, \qquad (5.69)$$

* بالنسبة لمعظم التطبيقات الاقتصادية، تعد قيمة $k = 3$ معقولة.

والذي يمثل الشكل العادي.

دع $\hat{a}_0$ ، $\hat{a}_1$ ، ... و $\hat{a}_{k_m}$ هي مقدرات المعلمات للمعادلة (5.69)، حينئذ فإن العلاقة المقدرة بين Y_t و X_t هي:

$$\hat{Y}_t = \hat{a}_0 + \hat{a}_1 X_1 + \cdots + \hat{a}_{k_m} X_t^{k_m}. \quad (5.70)$$

لأن مقدرنا $f(x_t)$ سيحُصل عليه بوساطة:

$$\hat{f}(X_t) = \hat{a}_0 + \hat{a}_1 X_1 + \cdots + \hat{a}_{k_m} X_t^{k_m}. \quad (5.71)$$

اعتبر الآن مشكلة ما إذا كان المتغير Y_t يعتمد على X_t أم لا. عند النظرة الأولى قد يظهر أن هذا الافتراض يمكن اختباره ببساطة عن طريق اختبار الفرضيات، واحدا بعد الآخر، بأن $a_1=0$ ، $a_2=0$ ، ... ، $a_{k_m}=0$ في المعادلة (5.64). وسوف نخلص افتراضا إلى أن Y_t و X_t مرتبطان بقوة ببعضهما بعضا عند رفض أي من فرضيات العدم هذه. وعلى العكس، إذا قبلنا جميع فرضيات العدم هذه ستصبح النتيجة هي أن Y_t و X_t غير مرتبطتين بقوة ببعضهما بعضا.

ولسوء الحظ، لايمكن اختبار هذه الفرضيات المرتبطة بعلاقة بين Y_t و X_t بهذه الطريقة بسبب مايمكن أن يسمى «خدعة التجميع» "Fallcy of Composition". أي أن الفرضية في هذه الحال، ترتبط بأكثر من معلمة. بخاصة، أن الفرضية التي نرغب في اختبارها هي:

$$H_0: a_1 = a_2 = \cdots = a_{k_m} = 0. \quad (5.72)$$

ولقد طورنا في ملحق هذا الفصل طريقة لاختبار الفرضيات على شاكلة المعادلة (5.75)، ولكننا نشير عند هذه النقطة إلى أنه، إذا كان $a_1=0$ ، $a_2=0$ ، ... ، $a_{k_m}=0$ قد اختيرت واحدة تلو الآخرى عند مستوى معين من المعنوية، مثلا 5%، وقبلت فإن امكانية رفض الفرضية (5.72) لاتزال عند المستوى نفسه من المعنوية. وبمعنى آخر، فإن الفرضية التي ترتبط بأكثر من معلمة، مثلا k_m ، كما في المعادلة (5.72) لايمكن، عمومًا، اختبارها على نحو عام عند مستوى معنوية عن طريق تقسيمها إلى فرضيات عددها k_m كل منها يرتبط بمعلمة مفردة ثم اختبار هذه الفرضيات

على التوالي عند ذلك المستوى من المعنوية.

إن توسيع طريقة التقدير أعلاه للحالة التي يشتمل فيها نموذج الانحدار على متغيرات مستقلة اضافية سهل ومباشر. افترض مثلا أن نموذجا من الشكل:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + f(X_{3t}) + u_t, \tag{5.73}$$

حيث إن u_t، مرة أخرى، هي الخطأ العشوائي، وأننا نفترض، أيضا، أن الشكل رقمالمعين لــ$f(X_{3k})$ غير مؤكد. حينئذ، يترتب على منهجنا أعلاه أنه إذا كانت فرضياتنا المرتبطة بـ$f(X_{3t})$ توحي بــأن $k = k_m$ سيعطي تقريبا ملائما لمتعدد الحدود، فسوف نفترض أن:

$$f(X_{3t}) = a_0 + a_1 X_{3t} + \cdots + a_{k_m} X_{3t}^{k_m}. \tag{5.74}$$

وبالتعويض من المعادلة (5.74) في المعادلة (5.73) ينتج:

$$Y_t = A + b_1 X_{1t} + b_2 X_{2t} + a_1 X_{3t} + \cdots + a_{k_m} X_{3t}^{k_m} + u_t, \tag{5.75}$$

حيث إن $A = a_0 + b_0$، مرة أخرى بجعل:

$$Z_{it} = X_{3t}^{i}, \qquad i = 1, 2, \cdots, k_m, \tag{5.76}$$

فإنه يمكن كتابة المعادلة (5.75) على النحو:

$$Y_t = A + b_1 X_{1t} + b_2 X_{2t} + a_1 Z_{1t} + \cdots + a_{k_m} Z_{k_m t} + u_t, \tag{5.77}$$

والتي هي في الشكل العادي. ويترتب على ذلك أنه– وبافتراض امكانية تقريب متعدد الحدود – يمكننا الحصول على مقدرات غير متحيزة لـ $\hat{A}$، $\hat{b}_1$، $\hat{b}_2$، $\hat{a}_1$، ...، $\hat{a}_{k_m}$ للمعاملات في المعادلة (5.77).

لاحظ أننا، في هذه الحال، لن نقدر على الحصول على مقدرات منفصلة لـa_0 وb_0 لأن مجموع A هو الذي يظهر في المعادلة (5.77). وبخلاف الحال، المبسطة أعلاه، سنقدر على تقدير $f(X_{3t})$ حتى ثابت تجميعي additive constant، وبمعنى آخر سنقدر على تقدير الجانب المتغير لـ$f(X_{3t})$، مثلا $f_v(X_{3t})$:

$$\widehat{f_v(X_{3t})} = \hat{a}_1 X_{3t} + \cdots + \hat{a}_{k_m} X_{3t}^{k_m}. \tag{5.78}$$

وعادةما يكون الجزء المهم من $f(X_{3t})$ هو جزؤه المتغير، بسبب أن هذا الجزء يصف الطريقة التي يتغير بها Y_t مع X_{3t}.

في هذه المرحلة ينبغي أن يتضح أن بإمكاننا توسيع الطريقة أعلاه لتشتمل عمومًا، على حالة وجود أي عدد من المتغيرات المستقلة. وينبغي أن يكون واضحا أيضا أنه يمكن توسيعها لتشتمل على نموذج بأكثر من متغير مستقل يدخل بطريقة غير محددة على النموذج.*

توليفات من الأشكال الدالية

لتجميع مناقشتنا حول الأشكال الدالية، نؤكد أنه من المشروع تماما استخدام اشكال عديدة من التحويلات المختلفة في معادلة الانحدار نفسها. وفي الحقيقة، قد نلاحظ في بعض امثلتنا السابقة، أن واحدا أو أكثر من المتغيرات تظهر في شكل لوغاريتمي أو ربما في شكل عكسي، بينما لاتكون المتغيرات الأخرى قابلة لأي شكل من التحويلات بتاتا. وعلى سبيل مثال إضافي، اعتبر الشكل التالي الأكثر تعقيدا لعلاقة منحنى فليبس:

$$\dot{W}_t = b_0 + b_1\left(\frac{1}{R_t}\right) + b_2\pi_{(t-1)} + b_3\dot{P}_t + b_4\dot{P}_t^2 + b_5 \ln G_t + u_t, \qquad (5.79)$$

حيث:

$\dot{W}$ = التغير النسبي للأجور خلال الفترة t،

R_t = معدل البطالة في الفترة t،

$\pi_{(t-1)}$ = معدل الأرباح للمنشآت في الفترة t-1،

$\dot{P}_t$ = التغير النسبي للأسعار في الفترة t،

$\ln_{Gt}$ = اللوغارتم الطبيعي لمعدل النمو في قوة العمل في الفترة t،

u_t = الخطأ العشوائي.

---

* اعتبر على سبيل المثال، نموذجا يأخذ الشكل:

$$Y_t = a_0 + f_1(X_{1t}) + f_2(X_{2t}) + a_1 X_{3t} + u_t.$$

لاحظ أننا سنستخدم في المعادلة نفسها التحويل العكسي، التحويل اللوغاريتمي، علاقة مبطأة، وشكلا لمتعدد الحدود لواحد من المتغيرات المستقلة. والآن سنعتبر المعادلة (5.79) لكي نوضح طرق التحويل، ولكن، عند التطبيق، تكون هناك عادة «أسباباً» (فرضيات اقتصادية) وراء كل تحويل. فمثلا رأينا في الفصل الثالث أن علاقة خطية بينWوR ليست ملائمة تماما (بسبب أن R لايمكن أن يأخذ قيما سالبة)، وأن التحويل العكسي ربما يكون منطقيا. وبالنسبة لمعدل الأرباح$\pi_{(t-1)}$ قد نتوقع أن اتحادات العمال التي تقوم بالمفاوضات (حول معدلات الزيادة في الأجور) تستخدم معدل الأرباح (والذي يكون عادة متاحا للفترة المحاسبية السابقة) باعتباره أحد العناصر في عملية المفاوضة. فإذا كانت الأرباح في الفترة الأخيرة عالية بصورة غير عادية فإن اتحاد العمال قد يجد مبررا في المطالبة بزيادات مناظرة أكبر في الأجور. وبالمثل قد يتغير شكل متعدد الحدود لتغير السعر، لأن العلاقة بين تعديلات الأجور وتعديلات الأسعار قد تكون غير خطية وذلك إذا كانت التغيرات السعرية الطفيفة تحدث دون أن يلاحظها العمال بينما تكون التغيرات السعرية الكبيرة مؤشرا للمطالبات بالزيادات في الأجور (انظر الشكل رقم٥-١٠). وأخيرا، ولأن النمو في قوة العمل يرتبط بزيادة عرض العمل، فقد يكون له تأثير على المركز التفاوضي في قوة العمل. وفي هذا المجال، رأينا في الفصل الثالث أن المتغير الأكثر ملائمة هو الزيادة النسبية (وليست المطلقة)، لذلك فإن حجم التغير يقاس بالنسبة إلى المستوى الحالي لعرض العمل. وللتوضيح سنستخدم في المعادلة (5.79) التحويل اللوغاريتمي للمتغيرG_t. والنقطة المهمة وراء كل هذا هي ببساطة أن اختيار الشكل الدالي ليس أساسا عملية تجربة وخطأ، وانما ينبغي استخدام المعلومات النظرية الاجتماعية المتاحة لدينا في تحديد الشكل الأكثر ملاءمة لعلاقتنا الدالية.

وبالعودة إلى معادلة الأجور، يمكننا وضع المعادلة (5.79) في شكل خطي

باستخدام التحويلات التالية:*

$$Z_{1t}=\left(\frac{1}{R_t}\right),$$
$$Z_{2t}=\pi_{(t-1)},$$
$$Z_{3t}=\dot{P}_t,$$
$$Z_{4t}=\dot{P}_t^2,$$
$$Z_{5t}=\ln G_t.$$

وبالتعويض عن هذه التحويلات في المعادلة (5.79)، نحصل على الشكل الخطي لمعادلتنا على النحو التالي:

$$\dot{W}_t=b_0+b_1Z_{1t}+b_2Z_{2t}+b_3Z_{3t}+b_4Z_{4t}+b_5Z_{5t}+u_t. \tag{5.80}$$

ويمكننا ببساطة حساب قيمة Z's من القيم المشاهدة لكل من R ، π ، $\dot{p}$ و G ومن ثم باستخدام هذه القيم Z's وطريقتنا العادية لتقدير يمكن حساب $\hat{b}_0$ ، $\hat{b}_1$ ، $\hat{b}_2$ ، $\hat{b}_3$ ، $\hat{b}_4$ و $\hat{b}_5$.

كل هذا ينبغي أن يوضح مدى المرونة الموجودة في تقدير الإنحدار المتعدد. كثيرا ما أنتقد الاقتصاديون القياسيون خاصة في بداية استخدام تحليل الانحدار بسبب ازدياد درجة اعتمادهم على الشكل الخطي من العلاقات، ولكن ينبغي الآن أن تكون قادرا على رؤية أن الاستخدام المنطقي، والذكي، للتحويلات يجعل نموذج الانحدار المتعدد قادرا على تناول تشكيلة عريضة من الأشكال الدالية المعقدة.

(٥-٤) توضيح: الطلب على النقود

إحدى القضايا المركزية في اقتصاديات النقود هي نظرية الطلب على النقود وقياسها*. في الحقيقة فإن كثيرا مما يتعلق بهذه القضية، كالتأثير المحتمل للسياستين

* في الحقيقة، فإن التحويل $Z_{3t}=\dot{P}_t$ غير ضروري وقد وضع، فقط، لتحقيق الاتساق في وضع الرموز.

المالية والنقدية على النشاط الاقتصادي يعتمد على شكل دالة الطلب على النقود وعلى قيم معلماتها. وتفيد النظرية الاقتصادية أن الطلب على الأرصدة النقدية الحقيقية (أي الأرصدة النقدية الرسمية معدلة بالمستوى العام للأسعار وذلك لتثبيت قوتها الشرائية) يعتمد على ثلاثة أنواع على الأقل من المتغيرات: الدخل، معدل الفائدة على السندات (أو معدل العائد على الأصول المالية الأخرى) وربما صافي الثروة. باختصار فإنه مع زيادة دخل الأفراد يزداد طلبهم على الأرصدة النقدية لأغراض التبادل، ومع ارتفاع أسعار الفائدة ينخفض طلبهم على الأرصدة النقدية لأرتفاع تكلفة الفرصة البديلة الفعالة للاحتفاظ بالأرصدة النقدية (والتي، لم تكن تحقق فائدة في الأقل، إلا مؤخرا)، ومع ارتفاع ثروات الأفراد فإنهم سوف يميلون للاحتفاظ بقدر أكبر من الأرصدة النقدية كأحد أشكال الاحتفاظ بالثروة المتزايدة.

يمكننا تلخيص كل هذا في المعادلة:

$$M_d = f(Y, r, W), \tag{5.81}$$

حيث:

M_d = الطلب على الأرصدة النقدية،

Y = الدخل الحقيقي،

r = معدل الفائدة، وأخيرا،

W = الثروة (أو صافي الثروة).

حيث نتوقع أن يكون التأثير الجزئي لسعر الفائدة سالبا أما التأثير الجزئي لمتغيرات الدخل والثروة فنتوقع أن يكون موجبا.

قام عدد من الاقتصاديين ببحوث قياسية مكثفة بهدف تقدير معادلات تناظر المعادلة (5.81)، اعتمدت أعمالهم على مجموعة من الأشكال الدالية المتضمنة في

* لمعالجة أكثر تفصيلا يرجع إلى:

David E. Laidler, *The Demand for Money : Theories and Evidence*, 2nd ed., (New York: Dun-Donnelley, 1977).

عديد من التحويلات التي اعتبرناها في هذا الفصل. وتشتمل هذه، أيضا، على قيم مبطأة لبعض المتغيرات. وعلى سبيل المثال، نعرض هنا نتائج إحدى الدراسات (قام بها Martin Bronfenbernner and Thomas Mayer)*. ونقطة البداية لديهما هي وضع دالة الطلب على النقود على شكل حاصل ضرب كالتالي:

$$M_{dt} = b_0 Y_t^{b_1} r_t^{b_2} W_t^{b_3} M_{d(t-1)}^{b_4} e^{u_t}, \qquad (5.82)$$

حيث u_t هو الخطأ العشوائي، وأن جميع المتغيرات الأخرى قد عرفت من قبل في المعادلة (5.81)، وبأخذ اللوغاريتمات لجانبي المعادلة (5.82) استطاع الباحثان أن يحصلا على الشكل الخطي التالي:

$$\ln M_{dt} = \ln b_0 + b_1 \ln Y_t + b_2 \ln r_t + b_3 \ln W_t + b_4 \ln M_{d(t-1)} + u_t. \qquad (5.83)$$

ولأغراض التفسير، افترض أن u_t تحقق افتراضاتنا العادية كافة، ونعيد كتابة المعادلة (5.83) على النحو التالي:

$$(\ln M_{dt} - b_4 \ln M_{d(t-1)}) = B + b_1 \ln Y_t + b_2 \ln r_t + b_3 \ln W_t + u_t. \qquad (5.84)$$

حيث إن $B = \ln b_0$. في هذا الشكل يمكن أن يستخدم نموذج مثل المعادلة (5.83) مفسرًا للفرق بين القيمة الحالية للمتغير التابع($\ln M_{dt}$) وقيمته المبطأة والمضروبة في المعامل b_4. فمثلا افترض أننا نشعر بأن لوغاريتم الطلب على النقود يتقلب عشوائيا حول خط إتجاه يتزايد بمعدل ٣٪ مثلا. في هذه الحال، فإن توقعاتنا المتعلقة بقيم لمعاملات في المعادلة (5.83) أو المعادلة (5.84) هي $b_4 = 1.03$ و $B = b_1 = b_2 = b_3 = 0$. وسنعتقد بأن الطلب على النقود لايستجيب لمتغيرات الدخل وسعر الفائدة والثروة.

باستخدام بيانات سنوية للولايات المتحدة الأمريكية للفترة ١٩١٩ - ١٩٥٦م، قدر الباحثان المعادلة (3.83) بطريقة المربعات الصغرى وحصلا على:

---

* يرجع إلى:

Martin Bronfenbrenner and Thomas Mayer. "Liquidity Functions in the American Economy." *Econometrica*, 28 (1960), pp. 810-834.

$$\begin{aligned} \ln M &= 0.11 + 0.34\ln Y - 0.09\ln r - 0.12\ln W + 0.72\ln M_{t-1} \\ &\quad\;\; (.003)\;\; (.09) \qquad (.01) \qquad\; (.08) \qquad\;\; (.06) \\ R^2 &= 0.99, \end{aligned} \tag{5.85}$$

حيث تمثل الأرقام داخل الأقواس أسفل معاملات الانحرافات المعيارية المقدرة المناظرة. ولتوضيح النتائج، فإن قيم t المحسوبة والمناظرة لمتغيرات الدخل، معدل الفائدة، الثروة، والمتغير المبطأ للطلب على النقود هي على الترتيب 8.3، 9، 1.5 و12 تقريبا. وعندما نستخدم قاعدتنا التجريبية، تدل هذه النتائج على أنه إذا اعتبرنا أن فرضية العدم $H_0 : b_3=0$ مقابل الفرضية البديلة $H_1 : b_3 \neq 0$ عند مستوى معنوية 5% فإننا سنقبل فرضية العدم. وبالمقابل إذا افترضنا أي فرضية من فرضيات العدم الأخرى $H_0 : b_i = 0$ مقابل $H_i : b_i \neq 0$، حيث i=1، 2، 4 عند مستوى معنوية 5% فسنرفض فرضية العدم.

لاحظ أن فرضية العدم $H_0 : b_2 = 0$ لها أهمية خاصة عند الاقتصاديين النقديين. ذلك أن رفض هذه الفرضية سوف يمنحهم سببا للاعتقاد بأنه عند معدلات الفائدة الأعلى يحتفظ الأفراد بنسبة أصغر من ثرواتهم في شكل أرصدة نقدية حتى يستطيعون الاستفادة من العائد الأعلى المتاح من السندات ومن الأصول التمويلية الأخرى التي تعطي فائدة. أحد المتضمنات المهمة لهذه النتيجة هي أن السياسة المالية لها بعض التأثير على الطلب الكلي، فإذا لم يكن الطلب على النقود مستجيبا لسعر الفائدة فإن السياسة المالية سوف تحفز فقط تغيرات معوضة offsetting في الإنفاق الخاص بدون أي تأثير صافي على الإنفاق في الاقتصاد.*

* في هذه النقطة، انظر Laidler، المرجع نفسه، الفصل الثاني. باختصار، إذا كان الطلب على النقود لايتأثر بالتغيرات في معدل الفائدة، فإن المنحنى LM سيكون رأسياً. ونتيجة لذلك، فإن التخفيض الضريبي أو الزيادة في معدل الإنفاق الحكومي سيدفع معدلات الفائدة إلى أعلى ومن ثم يقلل الانفاق الخاص بكمية مساوية. في هذه الحال، فإن مستوى الناتج القومي الإجمالي يعتمد تماما على عرض النقود.

ملحق أ (A): قيود طرفية في إبطاء آلمون

يعالج هذا الملحق استخدام قيود طرفية مع منهج آلمون لتقدير الانحدار ذي فترات الإبطاء، وكما ذكرنا في هذا الفصل من قبل، فقد يشعر الاقتصادي بأنه يعلم ليس، فقط، نمط b's ولكنه يعلم، بالضبط، أيضا، قيمة أي منb_0 و b_k أو كليهما، وهذه القيمة عادة صفر. فإذا علمنا هذه القيم، فإنه ينبغي علينا محاولة إدخال هذه المعلومات في عملية تقديرنا لنموذج الانحدار.

افترض، على سبيل المثال، أننا نشعر بأن نمط b's يكون مشابها لذلك المنحنى (A) الموجود في الشكل رقم(٥ أ-١) حيث تتناقص في البداية قيم b's ثم تتزايد، وأخيرا تتناقص حتى تصبحb_k مساوية للصفر. يمكننا في هذه الحال، أن نضع قيدا طرفيا واحدا وهو $b_k = 0$. فإذا افترضنا من الناحية الأخرى أن نمط b's يأخذ شكل المنحنى(B) في الشكل رقم(٥ أ-١) فإنه في هذه الحال، يمكن أن يوجد قيدان طرفيان وهما $b_0 = b_k = 0$.

وعمومًا، يمكن لواضع النموذج فرض أي من القيود الطرفية السابقة أوكلها، وهذا المنهج هو تعميم مباشر لطرق التقدير المعروفة، افترض، مرة أخرى، وجود المعادلة (5.24)، وأن k=10 (أي أن معادلتنا تحتوي على عشر فترات إبطاء):

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \cdots + b_{10} X_{t-10} + u_t. \qquad (5A.1)$$

افترض أولا (متجاهلا القيود الطرفية) أن النمط المقترح b's يأخذ شكل متعدد الحدود من الدرجة الثالثة:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \alpha_3 i^3, \qquad i = 0, \cdots, 10. \qquad (5A.2)$$

حينئذ، بالتعويض من المعادلة (5A.2) لكل واحدة من الـ b في المعادلة (5A.1)، تنتج لنا المعادلة:

$$Y_t = a + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + \alpha_3 Z_{4t} + u_t. \qquad (5A.3)$$

حيث:

$$Z_{1t} = \sum_{i=0}^{10} X_{t-i}, \qquad Z_{2t} = \sum_{i=1}^{10} i X_{t-i},$$

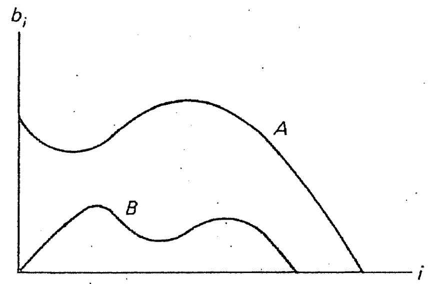
$$Z_{3t} = \sum_{i=0}^{10} i^2 X_{t-i}, \qquad Z_{4t} = \sum_{i=1}^{10} i^3 X_{t-i},$$

دعنا الآن نفرض قيدنا الطرفي، وبالتحديد افترض أننا نعتقد أن$b_{10} = 0$ حينئذ فمن المعادلة (5A.2)، يكون لدينا:

$$b_{10} = \alpha_0 + 10\alpha_1 + 100\alpha_2 + 1000\alpha_3 = 0. \qquad (5A.4)$$

ومن المعادلة (5A-4) يتضح أنه إذا فرضنا الشرط$b_{10} = 0$ فإنه ينبغي أن يكون لدينا:

$$\alpha_0 = -10\alpha_1 - 100\alpha_2 - 1000\alpha_3. \qquad (5A.5)$$



شكل (٥ أ-١)

وهذا يعني أن القيد$b_{10} = 0$ يتضمن قيدا على العلاقة بين α's، ويكون هذا، أساسا، هو حلنا للنموذج. ولكي نوضح ذلك، دعنا نعود إلى (5A.3) ونعوض عنα_0 من المعادلة (5A.5)، وهذا يعطينا:

$$Y_t = a + \alpha_1 Q_{1t} + \alpha_2 Q_{2t} + \alpha_3 Q_{3t} + u_t, \qquad (5A.6)$$

حيث:

$$Q_{1t} = Z_{2t} - 10Z_{1t},$$
$$Q_{2t} = Z_{3t} - 100Z_{1t},.$$

وأيضا:

$$Q_{3t} = Z_{4t} - 1000Z_{1t}.$$

وتعد المعادلة (5A.6) الشكل النمطي، وهكذا، فإنه يمكن تقدير كل من$\hat{\alpha}_1$، $\hat{\alpha}_2$ و$\hat{\alpha}_3$ باستخدام طرق تقدير الانحدار المتعدد النمطية، افترض أن $\hat{\alpha}_1$، $\hat{\alpha}_2$ و $\hat{\alpha}_3$هي

مقدراتنا، لذا يكون مقدرنا لـα_0 من المعادلة (5A.5) هو:

$$\alpha_0 = -10\hat{\alpha}_1 - 100\hat{\alpha}_2 - 1000\hat{\alpha}_3. \qquad (5A.7)$$

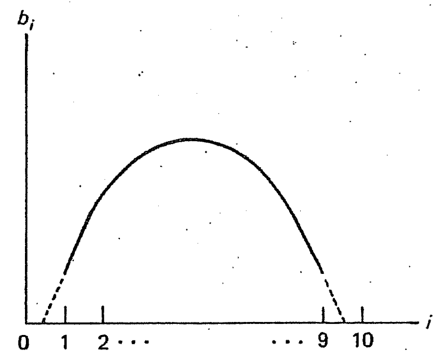
وأخيرا يمكن اشتقاق مقدراتنا b's من المعادلة (5A.2) على النحو التالي:

$$\begin{aligned}
\hat{b}_0 &= \hat{\alpha}_0 \\
\hat{b}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 \\
&\vdots \\
\hat{b}_9 &= \hat{\alpha}_0 + 9\hat{\alpha}_1 + 81\hat{\alpha}_2 + 729\hat{\alpha}_3 \\
\hat{b}_{10} &= 0.
\end{aligned} \qquad (5A.8)$$

وباختصار، فإن فرض القيد الطرفي$b_{10} = 0$ قد مكننا من الإحلال محلα_0 ومن ثم، إسقاط هذا المعامل من نموذج آلمون للانحدار (5A.3). ونترك للقارئ على سبيل التدريب، أن يثبت أن فرض القيد$b_0 = 0$ و$b_{10} = 0$ سينتهي بنا إلى إحلال تعبيرات فيα_2, α_3 محل كلا منα_0 وα_1 في نموذج الانحدار المتعدد (5A.3)، وهكذا فإن كل منα_0 وα_1 سوف يختفي من المعادلة التي نقدرها فعلا.

وقد ذكرنا في متن هذا الفصل أن استخدام هذه الطريقة غير المباشرة في فرض قيود طرفية سوف يعطي مقدرات غير متحيزة وذات تباينات أصغر من تلك الناجمة عن استخدام الطريقة المباشرة في التقدير (والتي تتضمن اسقاط كل منX_t وX_{t-k} من التحليل). ويكون ذلك صحيحا فقط في ظل افتراض محدد، وهو أن المعلمات الطرفية تقع على نفس المنحنى لمتعدد الحدود، كما هو الحال بالنسبة للمعلمات الأخرى غير الصفرية. فعلى سبيل المثال، إذا كانت القيود الطرفية هي$b_0 = b_{10} = 0$ فإنه ينبغي علينا أن نفترض أن قيمة متعدد الحدود تساوي الصفر عندما تكون $i=0$، $i=10$ وهذا مافعلناه بالنسبة لـ$b_{10} = 10$ في مثالنا السابق (5A.4).

ولكن قد لايتحقق هذا الافتراض، عند التطبيق. افترض، على سبيل المثال، أن $k=10$ وأن$b_0 = b_{10} = 0$، افترض أيضا، أن المعلمات غير الصفرية$b_1, ..., b_a$ تقع جميعها على متعدد للحدود من الدرجة الثانية مثل الموجود في الشكل رقم (٥أ-٢).



شكل (٥أ-٢)

في هذه الحال، يتضمن افتراض أن قيمة متعدد الحدود تساوي الصفر لكل من i=0، i=10 أنه، إذا تخيلنا متعدد الحدود سوف يقطع المحور الأفقي عند i=0 و i=10 . ويظهر شكل (٥أ-٢) أن ذلك ليس صحيحا بالضرورة. حيث يظهر الشكل رقمأن متعدد الحدود يقطع المحور الأفقي (كما يظهر في الشكل الخطوط المتقطعة) في مكان مابين i=0 و i=1 وأيضا بين i=9، i=10 وهكذا، إذا ما استخدمنا الطريقة السابقة في التقدير (5A.4) فإننا سوف ننتهي بفرض مجموعة من القيود الطرفية غير الصحيحة. وتكون النتيجة هي أن المقدرات الناتجة متحيزة. وباختصار لاينبغي علينا أن نفرض قيودا طرفية إلا إذا تبين من التحليل المتعمق صحتها. ولهذا السبب فإن الطريقة المباشرة في معالجة المعلومات المرتبطة بقيم المعلمات الطرفية (عن طريق، إسقاط X_t مثلا، من (5A.1) إذا كان يعتقد بأن معاملها b_0 يساوي الصفر) تبدو أكثر نجاحا.

ملحق ب (B):

اختبار الفرضيات المتضمنة لأكثر من معلمة انحدار

افترض أنه يوجد نموذج الانحدار التالي:

$$Y_t = a_0 + a_1 X_{1t} + \cdots + a_k X_{kt} + u_t, \qquad t = 1, \cdots, n, \qquad (5B.1)$$

حيث نفترض أن المتغيرات المستقلة X_1, ..., X_k، وأيضا، الأخطاء، العشوائية تحقق

جميع الفروض التقليدية لنموذج الانحدار، ومنها، بالطبع، افتراض أن الأخطاء العشوائية تكون موزعة توزيعا طبيعيا.

يرغب الاقتصاديون، عادة، في اختبار فرضيات النماذج المشابهة للنموذج (5B.1) التي تتضمن أكثر من معلمة واحدة، وتأتي هذه الفرضيات في أحد الشكلين التاليين: الأول منهما يرتبط بالقيود الخطية على المعاملات في (5B.1) وأحد الأمثلة لهذه الفرضيات هو:

$$\begin{aligned} a_1 &= a_2, \\ a_3 &= 2a_5, \\ a_0 + a_1 + a_2 + \cdots + a_k &= 1. \end{aligned} \tag{5B.2}$$

ويرتبط الشكل الثاني منهما بمعنوية مجموعة من المتغيرات المستقلة، افترض على سبيل المثال، أننا نريد اختبار الفرضية بأن Y_t لاتعتمد على أي من X_1 , X_2 أو X_3 في (5B.1) حينئذ، نكون مهتمين باختبار الفرضية:

$$a_1 = a_2 = a_3 = 0. \tag{5B.3}$$

وفي كلتا هاتين الحالتين السابقتين، أخذت الفرضيات الموضوعة في (5B.2) و (5B.3) على أنها فرضيات العدم، أما الفرضيات البديلة فقد أخذت على أنها مكملة لفرضية العدم (not H_0). على سبيل المثال، فإن الفرضية البديلة لفرضية العدم الموجودة في (5B.3) فرضية وجود واحد، في الأقل، من المعلمات a_1، a_2 و a_3 لايساوي الصفر. تكون الفرضية البديلة للفرضية الموجودة في (5B.2) هو:

$$\begin{aligned} a_1 &\neq a_2, \\ a_3 &\neq 2a_5, \\ a_0 + a_1 + a_2 + \cdots + a_k &\neq 1. \end{aligned} \tag{5B.4}$$

ويوجد، لحسن الحظ، منهج مباشر لاختبار مثل هذه الفرضية، ويمكن توضيح هذا المنهج في الخطوات التالية:

الخطوة الأولى: ندخل فرضية العدم موضع الاهتمام في نموذج الانحدار،

مثلا، إذا كانت الفرضية هـي أن $a_1 = a_2$، فإننا سوف نعيد كتابة نموذج الانحـدار الموجود في (5B.1) على النحو التالي:

$$Y_t = a_0 + a_1(X_{1t} + X_{2t}) + a_3 X_{3t} + \cdots + a_k X_{kt} + u_t. \tag{5B.5}$$

وعلى سبيل المثال، آخر، إذا كانت فرضية العدم هي $a_1 = a_2 = a_3 = 0$، فإننا نعيد كتابة النموذج (5B.1) على النحو التالي:

$$Y_t = a_0 + a_4 X_{4t} + \cdots + a_k X_{kt} + u_t. \tag{5B.6}$$

وأخيرا، وعلى سبيل توضيح ثالث، إذا كانت فرضية العدم $a_1 + 2a_2 + 5a_3 = 10$، فإنه سيكون لدينا $a_1 = 10 - 2a_2 - 5a_3$، ولذا يصبح النموذج:

$$\begin{aligned} Y_t = a_0 + 10X_{1t} + a_2(X_{2t} - 2X_{1t}) + a_3(X_{3t} - 5X_{1t}) \\ + a_4 X_{4t} + \cdots + a_k X_{kt} + u_t. \end{aligned} \tag{5B.7}$$

والذي يمكننا إعادة كتابته على النحو التالي:

$$\begin{aligned} (Y_t - 10X_{1t}) = a_0 + a_2(X_{2t} - 2X_{1t}) + a_3(X_{3t} - 5X_{1t}) \\ + a_4 X_{4t} + \cdots + a_k X_{kt} + u_t. \end{aligned} \tag{5B.8}$$

الخطوة الثانيـة: تقدير الشكل المقيد من نموذج الانحدار السابق وحسـاب مجموع مربعات الخطأ ESS_R.*

الخطوة الثالثة: تقدير الشكل الأولى (غير المقيد) لنموذج الانحـدار (5B.1) وحساب مجموع مربعات الخطأ ESS_U.

الخطوة الرابعة: تحديد الفرق في عدد المعلمات بين نموذجي الانحدار المقيد وغير المقيد، فإذا رمزنا لهذا الفرق بأنه d، تكون الفرضية المناظرة لـ $a_1 = a_2$، على سبيل المثال، فـي (5B.2) هي d=1. وعلى سبيل مثال، آخر نجد أن الفرضيـة فـي (5B.3) تتضمن أن d=3.

الخطوة الخامسة: نحسب النسبة التالية:

* تذكر أننا عرفنا من الفصلين الثاني والرابع أن $ESS = \Sigma(Y_t - \hat{Y}_t)^2$

$$\frac{(ESS_R - ESS_U)/d}{ESS_U/(n-k-1)}, \tag{5B.9}$$

حيث إن n هو عدد المشاهدات و (k+1) هو عدد المعلمات في النموذج الأصلي (غير المقيد)، ونقبل فرضية العدم أو نرفضها على أساس حجم هذه النسبة الموجودة في (5B.9) افترض، على سبيل المثال، أن فرضية العدم موضع الاهتمام غير صحيحة. حينئذ، سيكون نموذج الانحدار المقيد مبنيا على افتراض خاطئ ومن ثم، لن يكون محددا تحديدا دقيقا. ويكون النموذج الأصلي، في هذه الحالة، صحيحا. ويوضح ذلك أن كلا من ESS_R وESS_U سيختلفان عن بعضهما بعضا، وفي حالتنا هـذه، نتوقع بتحديد أكثر، أن يكون $ESS_U < ESS_R$.

ومن ناحية أخرى، افترض أن فرضية العدم صحيحة، حينئذ، سيكون نموذج الانحدار المقيد محددا تحديدا صحيحا. ولكن نموذج الانحدار غير المقيد سيكون، بدوره محددا بطريقة صحيحة أيضا. وعلى الرغم من أن ذلك قد يبدو غريبا، إلا أنه ليس من الصعب أن ندرك السبب. ذلك أنه عند التحديد الكامل للنموذج يكون الافتراض الوحيد الذي نختبره هو أن معلمات النموذج ثوابت. ومن الواضح أن هذا الافتراض سوف يتحقق سواء كانت فرضية العدم في (5B.2) أو (5B.3) صحيحة أم لا. ذلك أن شرط $a_1 = a_2$ لايخالف أي من الافتراضات المرتبطة بنموذج الانحدار كما في (5B.1). و، بالمثل، ولأن الصفر هو مقدار ثابت، فإنه إذا كانت فرضية العدم هي$(a_2 = a_3 = 0)$ فإن هذا، مرة أخرى، لن يؤدي إلى مخالفة أي من الافتراضات في (5B.1).

وفي فصل لاحق من هذا الكتاب، سوف نرى أنه إذا كانت الفرضية موضع الاهتمام صحيحة، فإن تطبيق نموذج الانحدار المقيد سوف يحقق فوائد، ترتبط بخصائص المقدرات. ولكننا، في المرحلة الحالية نحتاج أن نلاحظ، فقط، أنه إذا كان كل من الاشكال المقيدة وغير المقيدة في نموذج الانحدار قد حدد تحديدا صحيحا فإننا سوف نتوقع أن يكون الفرق بين ESS_R وESS_U صغيرا. وكل ذلك يمكن تلخيصه عن طريق ملاحظة أن القيم الكبيرة للنسبة الموجودة في (5B.9) توحي بأن

فرضية العدم تكون غير صحيحة بينما توحي القيم الصغيرة لها بصحة فرضية العدم.

وباستخدام اللغة الإحصائية يمكن أن نبين* أنه، إذا أعتبر فرضية العدم صحيحة فإن النسبة (5B.9) تأخذ شكل متغير F مع درجات حرية d و (n-k-1) أو باختصار $F_{d,n-k-1}$.

الخطوة السادسة: لما كانت القيم الصغيرة من هذه النسبة مرتبطة بقبول فرضية العدم فإننا نتوقع قبول H_0 عند مستوى معنوية 5% إذا كانت:

$$\frac{(ESS_R - ESS_U)/d}{ESS_U/(n-k-1)} < F^{0.95}_{d,n-k-1}, \qquad (5B.10)$$

حيث إن $F_{d,n-k-1}$ تكون:

$$\text{Prob}(F_{d,n-k-1} < F^{0.95}_{d,n-k-1}) = 0.95. \qquad (5B.11)$$

ويمكننا الحصول على $F_{d,n-k-1}$ من أي جدول من جداول توزيع F. ويوجد أحد هذه الجداول في نهاية الكتاب [الجدول الإحصائي رقم (٣)].

ولتوضيح الخطوات السابقة، دعنا نأخذ مثالا، افترض أن لدينا الافتراض التالي $a_1 = a_2, = a_3 = 0$. افترض أيضا، أن n=49 و k=8، نجد من الجدول الإحصائي رقم (3) أن: $F^{0.95}_{3,40} = 2.84$، افترض أننا حددنا ESS_R وأن $ESS_v = 20$، تكون نسبتنا حينئذ هي:

$$\frac{(35-20)/3}{20/40} = \frac{15/3}{1/2} = 10 > 2.84.$$

وهذا يعني أننا سنرفض عند مستوى معنوية 5% الفرضية: $a_1 = a_2, = a_3 = 0$.

---

* يرجع إلى: Arther Goldberger. *Econometric Theory* (New York, 1964), pp. 173-177.

أسئلـــة

١- افترض أن دالة للإنتاج تأخذ الـشـكــل الـتــالــي: $Q_t = (1/A)\, L_t^a\, K_t^b\, e^{u_t}$، حيث Q_t، L_t و K_t هي الإنتاج والعمل ورأس المال في الزمن t على التوالي، وأن u_t هي الأخطاء العشوائية المناظرة. افتـرض، أيـضـا، أن $E(u_t)=0$، $E(u_t^2)=\sigma^2$ و u_t مستقلة عن L_t و K_t. اقترح إحدى الطرق لتقدير A، a و b.

٢- افترض أن حجم الاستثمار الخاص في سنة معينة يعتمد على معدل الفائدة وعلى الحزب السياسي الذي ينتمي إليه الرئيس، بمعنى أن الاستثمار يكـون أكثر ارتفاعا في حالة ما إذا كان الرئيس ينتمي إلى الحزب الجمهوري بدلا من الحزب الديمقراطي. كون نموذجا يأخذ البيانات في شكل سلسلة زمنية على افتراض أن هناك حزبين فقط.

٣- افترض النموذج التالي للانحدار:

$$C_t = a_0 + a_1 F_1 Y_t + a_2 Y_t^{1/2} + a_3 (1/A_t) + u_t,$$

حيث C_t: الإنفاق الاستهلاكي للأسـرة t، Y_t دخل تلك الأسـرة، F_t حجم الأسرة و A_t حجم الأصول السائلة التي تمتلكها الأسرة، حول هذا النموذج إلى نموذج خطي.

٤- افرض أنك؛ تريد تقدير دالة الاستهلاك الخطية البسيطة التالية $C_t = a + bY_t + u_t$ لعدد n من الأفراد. كيف يمكنك أن تأخذ في الحسبان الانتقال في يالدالة بين المستهلكين في الحضر والمستهلكين في الريف، إذا كان الحد الثابت من الدالة يتأثر بموقع الإقامة للفرد.

٥- افترض أن الإنفاق الاستثماري لإحدى المنشآت يعتمد على معدل الفـائـدة ومعدل الأرباح وأخيرا على معدل التغير في المبيعات باعتباره مؤشرا علـى التوقعات:

(أ) كون نموذج الانحدار المقابل.

(ب) افترض أنه، خلال فترة العينة، كانت أرباح هذه المنشأة تعادل خلال

مختلـف الفترات الزمنية ١٥٪. ناقش مشاكل التقدير التي تنتج عن ذلك.

٦- افترض أن معدل الاستثمار في فترة معينة t، I_t يعتمد على معدل الفائدة في تلك الفترة r_t، وعلى المبيعات في تلك الفترة، وسبع قيم ذات فترات إبطاء لمعدل المبيعات S_t، افترض، أيضا، أن الأوزان المناظرة لفترات الإبطاء هذه تتزايد في البداية حتى تصل إلى ذروتها ثم تتناقص بعد ذلك.

(أ) كون شكلا غير مقيد لهذا النموذج

(ب) كون شكل آلمون لهذا النموذج

(جـ) اكتب المعادلات الطبيعية لشكل آلمون السابق.

٧- افترض أنه يوجد لدينا نموذج الانحدار المتعدد التالي:

$$Y_t = a + b_0 X_t + \cdots + b_6 X_{t-6} + \varepsilon_t .$$

افترض أننا استخدمنا منهج آلمون مع متعدد حدود من الدرجة الرابعة لتقدير معلمات هذا النموذج. افترض أخيرا أن النتائج كانت، على النحو التالي:

$$\hat{\alpha}_1 = 3, \quad \hat{\alpha}_2 = 5, \quad \hat{\alpha}_3 = 4, \quad \hat{\alpha}_4 = -10.$$

(أ) مإذا يكون تقديرنا لـ b_2 ؟

(ب) افترض أننا جعلنا $b_0 + b_1 + \dots + b_6 = 1$. عبر عن هذه المعلومة في شكل قيد على α's في تقريب متعدد الحدود لآلمون.

٨- يقال، أحيانا، إن النتائج التطبيقية لنموذج الانحدار يجب ألا تستـخـدم للتنبؤ بالحوادث التي تقع بعيدا جدا عن مجال التجربة. ناقش هذه العبارة (**مساعدة للحل**: راجع الافتراضات المرتبطة بتقريب متعددات الحدود).

٩- حول نموذج كويك التالي إلى شكل أبسط:

$$Y_t = a_0 + a_1 X_t + b_0 Z_t + b_1 Z_{t-1} + \cdots + u_t ,$$

حيث

$$b_i = b_0 \lambda^i , \qquad i = 1, 2, \cdots$$

١٠- اعتبر نموذج الإبطاء التالي لآلمون:

$$Y_t = b + b_0 X_t + b_1 X_{t-1} + \cdots + b_{10} X_{t-10} + u_t$$

حيث نفترض التقريب التالي من الدرجة الثانية:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2, \qquad i = 0, 1, \cdots, 10.$$

افترض أن $b_5 = 3$. المطلوب هو أن تشتق القيود المناظرة على α's، ومن ثم، الشكل المقيد لنموذج الانحدار.

١١- افترض أن الإنفاق الاستهلاكي C_t يعتمد على الدخل Y_t وعلى معدل الفائدة r_t إذا كانت r_t تزيد عن 0.05، أما إذا كانت r_t لاتزيد عن 0.05 فإن C_t تعتمد فقط على Y_t. كون هذه العلاقة في شكل نموذج للانحدار.

١٢- حول النموذج التالي إلى نموذج انحدار خطي متعدد، ثم اكتب المعادلات الطبيعية له:

$$\log Y_t = a_0 + a_1 e^{X_{1t}} + a_2 \left(\frac{1}{1 + X_{1t} X_{2t}} \right) + u_t.$$

الفصل السادس

مشاكل في تحليل الانحدار

تناولنا في الفصول السابقة طرق تدير العلاقة بين مجموعة من المتغيرات، وتعلمنا كيف يمكننا استخدام هذه المعلومات في اختبار الفرضيات وفي توقع تأثير تغير أحد المتغيرات على الآخر. تعتمد طرق التقدير هذه على عدد من الفروض التي ناقشناها في معالجتنا لنموذج الانحدار. ولكن، عند التطبيق، يجد الاقتصاديون أن من المؤكد (أو المرجح في الأقل في بعض الحالات الأخرى) أن واحدا أو أكثر من هذه الفروض لن يتحقق. إما بسبب طبيعة العلاقة الدالية التي تجمع بين متغيرات النموذج، أو نتيجة لوجود صعوبات كبيرة تنشأ من مجموعة معينة من القيم المشاهدة للمتغيرات. وتتجه الآن معظم الأعمال الحديثة في مجال الاقتصاد القياسي نحو تطوير طرق تقدير معدلة وبنائها لتعالج مثل هذه المشكلات.

وسيكون ذلك هو موضوع الفصلين الأخيرين من هذا الكتاب. سنتعرض هنا للحالات التي تسبب مخالفة افتراضات نموذج الانحدار، أو التي تتسبب - في الأقل - في إيجاد صعوبات تمنع الاستفادة منها أو تقلل فاعليتها، وبعد ذلك سنناقش الكيفية التي يمكن أن نعدل بها طرق تقديرنا لنأخذ في الاعتبار هذه المشاكل.

(٦-١) تعدد العلاقات الخطية

ناقشنا، بالفعل، في أماكن عديدة، مشكلة تعدد العلاقات الخطية multicollinearity وبلغة فنية تنشأ هذه المشكلة عندما يكون واحد، في الأقل، من المتغيرات المستقلة توليفة خطية من المتغيرات الأخرى. وينتج عن ذلك وجود عدد قليل جدا من المعادلات الطبيعية المستقلة، ومن ثم، عدم امكانية اشتقاق مقدرات للمعاملات الموجودة بالنموذج كافة. وعلى سبيل مراجعة مختصرة افترض أن:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t. \qquad (6.1)$$

حيث تكون قيم كل من X_1 وX_2 دائما منطبقة على بعضها بعضا بمعنى أن لدينا:

$$X_{1t} = X_{2t} \qquad , \qquad t = 1, 2, \cdots, n.$$

وهذا يعني أن أي تحرك من فترة لأخرى في المتغير X_1 يناظر، تماماً، تحرك مماثل في X_2. فإذا كان ذلك صحيحا فإنه لايمكننا أن نعزل تأثير X_1 على Y عن تأثير X_2 على Y. لكن، تذكر أنه مازال من الممكن تقدير الأثر المشترك لكل من X_1 وX_2 علىY أي أنه، بإحلالX_{1t} محلX_{2t} في المعادلة (6.1)، نستطيع تقدير قيمةb_3 حيث إنها تساوي($b_1 + b_2$) في المعادلة (6.2).*

$$Y_t = b_0 + (b_1 + b_2) X_{1t} + u_t = b_0 + b_3 X_{1t} + u_t. \qquad (6.2)$$

وإذا ما استمر وجود تعدد العلاقات الخطية بينX_{1t} وX_{2t} في فترات خارج الفترة التي تغطيها عينتنا فإن الصيغة المقدرة للمعادلة (6.2) تصبح ملائمة للتنبؤ بالقيم المستقبلية للمتغير التابع Y_t.

ويطلق على حالة تعدد العلاقات الخطية هذه تعدد العلاقات الخطية التام، حيث يكون واحدا أو أكثر من المتغيرات المستقلة توليفة خطية تامة من المتغيرات الأخرى. وتنتج هذه الحالة عن الطريقة التي تكون بها معادلة الانحدار، وتشبه هذه الحالة تلك التي ناقشناها في الفصل السابق حيث استخدمت المتغيرات الصورية

---

* يمكننا الآن رياضيا أن نقدر المعادلة (6.2) ذلك لأننا قللنا عدد المعلمات التي ينبغي أن نقدرها بمقدار معلمة واحدة، ولذلك، فإن عدد المعلمات أصبح يساوي عدد المعادلات الطبيعية المستقلة.

لكل الفصول المناخية الأربعة (بدلا من ثلاثة فقط). وعلى سبيل مثال آخر، افترض أننا وضعنا دالة الاستهلاك في الشكل:

$$C_t = b_0 + b_1 X_{dt} + b_2 Y_{d(t-1)} + b_3 (\Delta Y_{dt}) + u_t, \tag{6.3}$$

حيث Y_{dt} تعكس تأثير الدخل الجاري على الاستهلاك $Y_{d(t-1)}$ تشير إلى تأثير مستويات الدخل الماضية أو تأثير العادة، بينما تعكس ΔY_{dt} و $[\Delta Y_{dt} = Y_{dt} - Y_{d(t-1)}]$، أثر التوقع الناجم عن التغيرات في مستويات الدخل، ولما كانت ΔY_{dt} توليفة خطية من كل من Y_{dt} و $Y_{d(t-1)}$ فلن يكون باستطاعتنا تقدير كل المعاملات في المعادلة (6.3). وسوف نترك للقارئ أن يضع المعادلة (6.3) في صيغة يمكن تقديرها واستخدامها في التنبؤ بقيم C_t.

هذه إذن حالات من تعدد العلاقات الخطية التام. ولكن هذه المشكلة تظهر أيضا، بدرجات مختلفة، وفي العادة تأتي بدرجة أقل من تعدد العلاقات الخطية التام، وهذه الأخيرة هي التي تسبب معظم المشاكل للباحثين. وتنشأ هذه المشكلة عندما ترتبط المتغيرات المستقلة ببعضها بعضا ارتباطًا قويًا، ولكن غير تام. افترض على سبيل المثال أننا نحاول تقدير دالة الطلب كما في المعادلة (6.4).

$$Q_t = b_0 + b_1 P_t + b_2 Y_t + u_t. \tag{6.4}$$

لمجموعة معينة من السلع، ولتكن الواردات، حيث نفترض أن الكمية المطلوبة (Q) تعتمد على مستوى الأسعار للسلع المنتجة محليا (p) وعلى مستوى دخل المستهلكين. ومن المعروف أنه إذا كانت الأسعار المحلية مرتفعة زاد الطلب على المنتجات الأجنبية التي تعد أرخص نسبياً، ولذا، نتوقع أن تكون b_1 موجبة. وبالمثل، فإنه إذا كانت دخول المستهلكين أكبر كان الطلب أكبر على السلع (بما فيها الواردات)، ولذا نتوقع أن تكون b_2 موجبة ايضاً. ونلاحظ بفحص البيانات المتاحة، أنه على مدى الفترات الزمنية التي يزداد فيها معدل التضخم المحلي، تزداد الواردات، وعلى نحو مشابه، تنمو الواردات أيضا، خلال فترات تزايد الدخل. والصعوبة هنا هي أن فترات تزايد الدخل هي، عموما، فترات التضخم المرتفع والعكس صحيح. أو بمعنى آخر، يوجد ارتباط موجب قوي (وإن لم يكن تاماً) بين p و Y.

هذا الارتباط القوي يجعل من الصعوبة بمكان عزل آثار p و Y على Q فعندما يكون هناك تزايد سريع في الواردات في الوقت نفسه الذي يزداد فيه كلٍ من الدخل والأسعار المحلية، يكون من الصعب تحديد الآثار النسبية لكل من التضخم والدخول الأعلى في حفز الزيادة في الواردات.

تعدد العلاقات الخطية غير التام: بعض النتائج المنطقية

كيف يعرف الباحث أن لديه مشكلة تعدد علاقات خطية خطيرة في نموذجه؟ كما أشرنا من قبل تظهر مشكلة تعدد العلاقات الخطية في درجات مختلفة وقد تسبب مشاكل عسيرة في بعض الحالات وقد لاتكون كذلك في حالات أخرى. وهناك على أي حال، بعض نتائج الانحدار التي يمكن تفسيرها، فقط، بدلالة الدرجات العالية من تعدد العلاقات الخطية غير التام: منها مثلا عندما يوجد معامل تحديد كبير R^2 مصاحب لتقديرات غير معنوية إحصائيا لمعاملات المتغيرات المستقلة. ويعني ذلك أن هناك متغيرات مستقلة معينة (أو في الأقل، واحد منها) تؤثر تأثيرا منتظما في المتغير التابع (كما يشار إليه بوساطة قيمة R^2 المرتفعة) ولكن لايمكننا معرفة أي منها هو المسئول عن ذلك التأثير.

وتظهر مشكلة تعدد العلاقات الخطية بدقة أكثر في وجود تباينات كبيرة لمقدرات المعاملات. ولما كان التباين الكبير يعني أن نسبة مئوية (95% مثلاً) لفترة ثقة معينة للمعلمة المناظرة ستكون عريضة نسبياً، فإن مدى كبير من القيم للمعلمة (وربما تتضمن هذه قيمة الصفر) ستكون متسقة مع فترتنا للثقة. ويعني ذلك – أنه حتى إذا كان المتغير المستقل المناظر له تأثير مهم على المتغير التابع، فإن مشكلة تعدد العلاقات الخطية قد تجعل من الصعب القيام بتقدير أثر ذلك المتغير بدقة. ومن ثم، تنخفض درجة الثقة في السياسات المقترحة بناء على هذه المقدرات.

وحتى نرى كيف تنتج تباينات كبيرة عن مشكلة تعدد العلاقات الخطية غير التام (ومن ثم، انحرافات معيارية كبيرة) لمقدراتنا، تذكر أننا بينا في الفصل الرابع أن تباين المقدر $\hat{b}_1$ هو:

$$\operatorname{var}(\hat{b}_i)=\frac{\sigma_u^2}{\sum \hat{v}_{it}^2} \tag{6.5}$$

حيث $\hat{v}_{it}$ هو الباقي في انحدار المتغير المستقل رقـم i^{th} (X_{it}) على جميع المتغيـرات المستقلة الأخرى في النموذج $\hat{v}_{it}=X_{it}-\hat{X}_{it}$ فإذا كانت هناك علاقة خطية وثيقة بين المتغيرات المستقلة، فإن $\hat{v}_{it}$ سيصبح صغيرا لأن القيمة المحسوبـة $\hat{X}_{it}$ سوف تكون قريبة من القيمة الفعلية للمتغير X_{it}. وينتج عن ذلك صغر المقام في المعادلة (6.5) مما يوجد تباينا كبيرا لـ $\hat{b}_i$.

ومن المهم أن نفسر بدقة مايعنيه كل هذا. فيجب أن نتذكر أن ذلك لايعني أن تقديرات المعاملات متحيزة، بل أنها تظل غير متحيزة، كما أثبتنا ذلـك فـي ملحق الفصل الرابع. غير أن مايؤدي إليه تعدد العلاقات الخطية هو عـدم دقـة مقدراتنا: حيث يصبح تباينها كبيرا، ومن ثم، لايمكن الاعتماد عليها اعتـمـادا كبيراً. وتكمن المشكلة - كما ذكرنا من قبل - في صعوبة تحديد تأثير كل متغير مستقل - بعيدا عن تأثير المتغيرات الأخرى على المتغير التابع.

تعليق إضافي

افترضنا، ضمنيا حتى الآن، أن جميع المتغيرات المستقلة بالنموذج مرتبطة ببعضها البعض ارتباطا قوياً. وقد لايكون الوضع كذلك بالضرورة. افترض مثلاً، أن نمـوذج الانحدار يحتوي على متغيرات مستقلة X_{1t}, X_{2t}, X_{3t}. افترض، أيضا، أن X_{1t} و X_{2t} مرتبطتان ببعضهما بعضا ارتباطا قويا (وإن كان هذا الارتباط ليس تاما) ولكن X_{3t} ليست مرتبطة نسبيا بكل من X_{1t} و X_{2t} . حينئذ فإن صيغة التباين في المعادلة (6.5) تفيد أن تباين المقدرات المناظرة للمعاملات X_{1t} و X_{2t} (ولتكن $\hat{a}_1$ و $\hat{a}_2$) ستكون كبيرة بينما يكون تباين X_{3t} $(\hat{a}_3)$ وهو معامل X_{3t} ليس كذلك بالضرورة. وبديهيا إذا كانـت X_{3t} ليست مرتبطـة بقوة بكل من X_{1t} و X_{2t} فإن $\hat{v}_{3t}^2$ يمكن أن تكون كبيرة، عمومًا طالما أن $\hat{X}_{3t}$ تمثل تنبؤا غير دقيق لـ X_{3t} (بمعنى أن X_{3t} لايمكن تفسيرها مرضيا بوساطة X_{1t} و X_{2t}).

والآن يمكننا أن نضع قاعدة قد تكون مفيدة في تفسير مشكلة تعدد العلاقات الخطية. ذلك أن المقدار $\Sigma\hat{v}_{it}^2$ والذي يظهر في مقام المعادلة (6.5) هو مجموع مربعات الخطأ الذي حصل عليه عن طريق تكوين انحدار لـX_{it} على جميع المتغيرات المستقلة الأخرى الموجودة بالنموذج. دعنا نرمز إلى مجموع مربعات الخطأ هذا بالرمز ESS_i وأن $TSS_i = \Sigma X_{it}^2$ وحينئذ فإن $TSS_i = RSS_i + ESS_i$ ومنها فإن $RSS_i = TSS_i - ESS_i$. ومن الواضح أن $ESS_i \geq 0$ وأن $TSS_i \geq 0$. ويمكن أن نثبت، أيضا، أن $RSS_i \geq 0$*، ومن ثم، فإن $TSS_i \geq RSS_i$ و $TSS_i \geq ESS_i$.

دعنا نرمز للنسبة RSS_i/TSS_i بالرمز r_i^2. حينئذ، فإن $0 \leq r_i^2 \leq 1$. ومن الواضح أنه كلما كانت قيم X_{it} مرتبطة إرتباطا قويا بقيم المتغيرات المستقلة الأخرى كانت ESS_i صغيرا، ومن ثم، تكون r_i^2 قريبة من الواحد الصحيح. وعلى العكس من ذلك، كلما كانت العلاقة بين X_{it} والمتغيرات المستقلة الأخرى ضعيفة، كانت ESS_i كبيرة، ومن ثم، تكون r_i^2 قريبة من الصفر. لذا يكون r_i^2 مشابها لمعامل التحديد المتعدد الذي يربط X_{it} بجميع المتغيرات المستقلة الأخرى في النموذج.

والآن دعنا نعبر عن مقام صيغة التباين بالمعادلة (6.5) صورة أخرى من خلال معرفة $ESS_i = TSS_i - RSS_i \equiv TSS_i\ (1-r_i^2)$ وينتج عن ذلك أن تباين $\hat{b}_i$ يمكن أن نعبر عنه على النحو التالي:

$$\operatorname{var}(\hat{b}_i) = \frac{\sigma_u^2}{TSS_i(1-r_i^2)} \equiv \frac{\sigma_u^2}{\sum X_{it}^2\,(1-r_i^2)} \qquad (6.6)$$

ويشير التعبير الموجود في المعادلة (6.6) بوضوح لتأثير تعدد العلاقات الخطية الجزئي على تباين أحد المقدرات. ويناظر غياب تعدد العلاقات الخطية الجزئي الحالة التي يكون فيها $r^2 = 0$. وتصبح المشكلة أكثر صعوبة وتؤدي إلى تباين أكبر كلما اقتربت r من الواحد الصحيح. لاحظ أخيرا أن r ليس من الضروري أن يتساوى في الكبر لكل من $i=1,...,k$.

---

* أنظر (4A.5) في الملحق للفصل الرابع ولاحظ أنه، بالنسبة للمتغير المستقل رقم k^{th} يكون $TSS_k' = \Sigma X_{kt}^2$، $ESS_k = \Sigma\hat{v}_{kt}^2$ وأن $RSS = \Sigma\hat{X}_{kt}^2$ وذلك لأن $\Sigma\hat{X}_{kt}\hat{v}_{kt} = 0$.

بعض الحلول

ليس من السهل حل مشكلة تعدد العلاقات الخطية، غير أنه يمكن للباحث، دائماً، أن يحاول زيادة دقة مقدراته (أي تخفيض تبايناتها) عن طريق زيادة عدد المشاهدات. وعلى سبيل المثال، فإنه، مهما صغرت قيم v_{it}^2 في المعادلة (6.5) فإنه من الواضح أن تباين $(\hat{b}_i)$ يتناقص مع زيادة حجم العينة. ولكن، من الواضح، أيضا، أن زيادة عدد مشاهدات العينة ليس ممكنا دائماً، وفي حالة كون مشكلة تعدد العلاقات الخطية خطيرة بما فيه الكفاية، فإن زيادة حجم العينة قد لايساعد كثيرا اللهم إلا إذا كانت هذه الزيادة كبيرة جداً.

ومن الطرق البديلة لعلاج مشكلة تعدد العلاقات الخطية إضافة معلومات يمكن استخدامها في تقدير قيم المعاملات الفردية. افترض، على سبيل المثال، أننا نرغب في تقدير دالة الإنتاج المبينة في المعادلة (6.7) لسلعة ما:

$$Q_t = AL_t^{\alpha} K_t^{\beta} e^{u_t}, \tag{6.7}$$

حيث ترمز Q_t الكمية المنتجة في الفترة t، L_t لساعات العمل، k_t لرأس المال، u_t للخطأ العشوائي و β, α, A هي المعلمات التي ينبغي تقديرها. تذكر أنه يمكن، عن طريق استخدام التحويل اللوغاريتمي وضع المعادلة رقم (6.7) في شكل يمكن تقديره على النحو التالي:

$$Q_t^* = A^* + \alpha L_t^* + \beta K_t^* + u_t. \tag{6.8}$$

حيث ترمز النجوم إلى لوغارتمات المتغيرات الأولية في المعادلة (6.7). افترض، للتوضيح، أنه توجد لدينا مشكلة تعدد علاقات خطية جزئي في العينة موضع البحث يأخذ شكل ارتباط قوي بين L و k في هذه الحال، يؤدي الارتباط القوي بين L وK (ضمن أشياء أخرى) إلى إيجاد تباينت كبيرة لمقدرات معلمات مرونة دالة الإنتاج α وβ .

والآن، دعنا نفترض أنه توجد دلائل قوية، بناء على المعلومات التي أمكن الحصول عليها من مصدر آخر، تشير إلى أن هذه الصناعة المنتجة لهذه السلعة

تخضع لقانون ثبات غلة الحجم. ومن مناقشتنا لدوال الإنتاج في الفصل الأخير، نجد أن هذا يعني أن: $(\beta+\alpha=1)$ ويمكننا الآن بوساطة هذه المعادلة أن نستبدل β بـ $(1-\alpha)$ في المعادلة (6.7) لكي نحصل على:

$$Q_t = AL_t^{\alpha} K_t^{(1-\alpha)}\, e^{u_t}. \qquad (6.9)$$

وبأخذ اللوغاريتمات، يمكننا أن نحصل على:

$$Q_t^* = A^* + \alpha L_t^* + (1-\alpha)K_t^* + u_t, \qquad (6.10)$$

حيث ترمز النجوم - مرة أخرى - إلى لوغارتمات المتغيرات الأولية. وبإعادة ترتيب مكونات المعادلة (6.10)، نحصل على:

$$Q_t^* - K_t^* = A^* + \alpha(L_t^* - K_t^*) + u_t \qquad (6.11)$$

أو $$Y_t^* = A^* + \alpha Z_t^* + u_t,$$

حيث: $Y_t^* = (Q_t^* - K_t^*)$ و $Z_t^* = (L_t^* - K_t^*)$.

وهكذا فإن معلوماتنا المسبقة مكنتنا من اختزال نموذجنا إلى نموذج يحتوي على متغير مستقل واحد Z_t^*. وينبغي أن يلاحظ القارئ أنه حتى ولو كانت L_t^* و K_t^* مرتبطين ببعضهما بعضا ارتباطا قويا فلن توجد، عمومًا، مشكلة ناتجة عن تعدد العلاقات الخطية في تقدير المعادلة (6.11)، فإذا افترض، مثلا أن $L_t^* = 4k_t^*$ فإنه لن توجد مشكلات تقدير بسبب تعدد العلاقات الخطية نتيجة لأن نموذجنا (6.11) سوف يختزل إلى:*

$$Y_t^* = A^* + \alpha(3K_t^*) + u_t. \qquad (6.12)$$

وباختصار فإن المعلومات الإضافية التي تبين أن الصناعة تخضع لقانون ثبات غلة الحجم قد مكنتنا من الحصول على مقدرات ذات تباينات أقل للمعلمات A و α.

* يبين هذا المثال أننا سنواجه مشكلة، فقط، إذا كانت $L_t^* - K_t^*$ تساوي مقدار ثابت. ويحدث ذلك عندما تكون L_t نسبة من K_t أي عندما يتحقق $L_t = dK_t$، حيث تمثل d مقدار ثابت. وفي هذه الحالة الخاصة ستكون $Z_t^* = \left(L_t^* - K_t^*\right)$ في المعادلة (6.11) مقدارا ثابتا وسوف لانتمكن من الحصول على مقدار لـ α. وفي هذا المجال تذكر من مناقشتنا لنموذج الانحدار البسيط في الفصل الثاني أن المتغير المستقل ينبغي أن يأخذ، في الأقل، قيمتين مختلفتين من أجل أن نتمكن من تقدير معاملات الانحدار.

ومقدرنا لـ β سوف يصبح، حينئذ:

$$\hat{\beta}=1-\hat{\alpha}. \quad (6.13)$$

تأثيره على التنبؤ

لاتتوافر في كثير من الأحيان معلومات إضافية يمكن استخدامها لتخفيف حدة مشكلة تعدد العلاقات الخطية، ويصطدم الباحث بمجموعة من تقديرات المعلمات لايمكن الاعتماد عليها. وربما نشعر بشئ من الراحة إذا علمنا أن المعادلة المقدرة ماتزال مقبولة لأغراض التنبؤ. وبأخذ مثال شاذ، افترض أن العلاقة التالية التي تظهر تعددا تاما للعلاقات الخطية وذلك بسبب الطريقة التي تعرف بها المتغيرات.

$$Y_t=b_0+b_1X_{1t}+b_2X_{2t}+u_t, \quad (6.14)$$

حيث $X_{1t}=3X_{2t}$. كما لوحظ من قبل وبسبب أن X_1 ترتبط بعلاقة خطية تامة مع X_2، فمن غير الممكن تقدير أي من b_2 و b_1. ولكن لغرض التنبؤ لانهتم بقيم b_2 و b_1 في ذاتها ولكننا نهتم بالقيمة المتوسطة لـ Y_t المناظرة لـ X_1 و X_2 أي:

$$\begin{aligned} Y_t^m &= b_0+b_1X_{1t}+b_2X_{2t} \\ &= b_0+(3b_1+b_2)X_{2t}. \end{aligned} \quad (6.15)$$

ونعرف من مناقشتنا الحالية أنه يمكننا تقدير القيمة المتوسطة لـ Y_t بوساطة المقدر:

$$\hat{Y}_t=\hat{b}_0+\hat{b}^*X_{2t}, \quad (6.16)$$

حيث $\hat{b}^*$ هي مقدرنا للمجموع $(3b_1+b_2)$. وهكذا، فعلى الرغم من أننا لانستطيع تقدير تأثير كل من X_1 و X_2 على Y_t إلا أنه يمكننا تكوين تنبؤات ترتبط بقيمة Y_t المناظرة لأي قيمة من X_{2t} طالما ظلت العلاقة $X_{1t}=3X_{2t}$ قائمة.* وتمتد هذه النتيجة لتشمل حالات تعدد العلاقات الخطية غير التام، غير أن إثبات ذلك يخرج عن

---

* ينبغي أن يكون واضحاً أن دقة التنبؤ أو جودته تعتمد على اعتبارين اثنين هما: (ا) دقة مقدراتنا للقيمة المتوسطة لـ Y_t (أي تباين $\hat{Y}_t$ باعتباره مقدرا لـ Y_t^m)، و ب) حجم تباين الخطأ العشوائي u_t. ولسوء الحظ، فإن المعادلات التي تصف دقة التنبؤ في إطار الانحدار المتعدد تتطلب معرفة مفاهيم إحصائية فوق المستوى المستخدم في هذا الكتاب.

نطاق هذا الكتاب. على الرغم من أن مقدراتنا للتأثيرات المنفصلة للمتغيرات المستقلة قد تكون لها تباينات كبيرة، فإن مقدرنا $\hat{Y}_t$ للأثر المشترك على Y_t، الذي يمكن أن يعبر عنه بوساطة القيمة المتوسطة لـ Y_t، أو Y_t^m، قد يكون له تباين صغير. ولما كانت تنبؤاتنا تتضمن تقدير متوسط Y_t، وأنه يمكننا تقدير المتوسط بدقة كبيرة، فإن مشكلة تعدد العلاقات الخطية قد لاتمثل مشكلة صعبة لأغراض التنبؤ.

(٦-٢) مشكلة الارتباط الذاتي

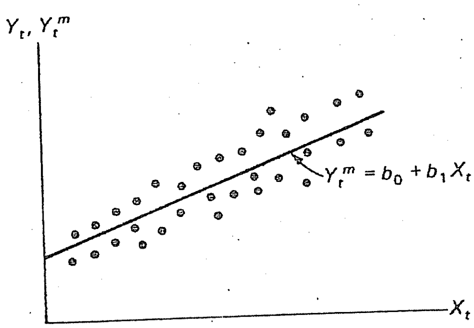
نعلم أن أحد الفروض الأساسية لنموذج الانحدار هو أن قيمة الخطأ العشوائي في إحدى الفترات الزمنية تكون مستقلة عن قيمته في أي فترة زمنية أخرى بحيث يمكن القول إن:

$$\operatorname{cov}(u_t, u_s) = 0 \quad , \quad t \neq s.$$

وتتضمن هذه المعادلة بالنسبة لقيم معطاة للمتغيرات المستقلة أن قيمة Y_t سوف تختلف عن قيمتها المتوسطة Y_t^m بمقدار مستقل عن حجم الخطأ العشوائي في أي فترة زمنية أخرى. وبالنسبة لنموذج الانحدار الذي يأخذ الشكل:

$$Y_t = b_0 + b_1 X_t + u_t,$$

نتوقع أن تأخذ خريطة الانتشار الصورة الموضحة في الشكل (٦-١) حيث تكون النقط المشاهدة «منتشرة عشوائيا» حول خط الانحدار.

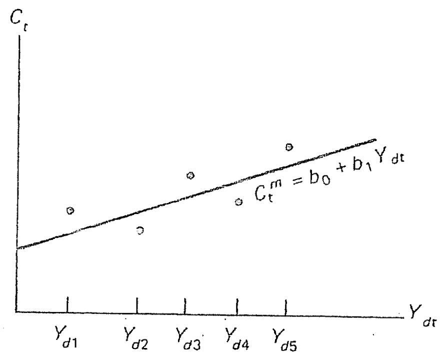


شكل رقم (٦-١)

افترض من ناحية أخرى، أن $\text{cov}(u_t, u_s) \neq 0$، لذلك، تكون القيم المتتابعة للخطأ العشوائي غير مستقلة عن بعضها بعضا. افترض، على سبيل المثال، أنه يمكن وصف السلوك الاستهلاكي لفرد ما بوساطة المعادلة:

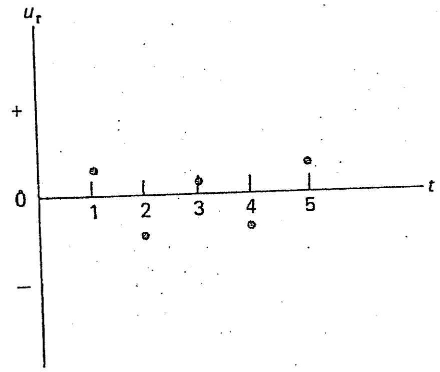
$$C_t = b_0 + b_1 Y_{dt} + u_t, \tag{6.17}$$

ولكن قيم u_t ليست مستقلة عن قيمها السابقة. فإذا كان هذا الفرد ينفق، مثلا، كثيرا، جدا في الفترة الأولى (ربما نتيجة زيارة غير متوقعة من بعض أصدقائه) بحيث تكون $u_1 > 0$ فسوف يحاول أن يعوض هذه الزيادة في الإنفاق في الفترة التي تليها عن طريق إنفاق قدر أقل من المعتاد، ومن ثم، نتوقع أن تكون $u_2 < 0$. لاحظ أن ذلك يتضمن بعمومية أكثر أن u_t لها ارتباط سالب مع u_{t+1}. فإذا كان مستوى الدخل يزداد في الفترات المتعاقبة فإن مثل هذا التصاحب السالب بين القيم المتتابعة للخطأ العشوائي يتوقع أن يولد خريطة انتشار تشبه، إلى حد ما، الشكل رقم (٦-٢).



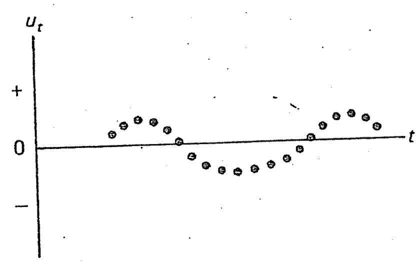
الشكل رقم (٦-٢)

فإذا أردنا أن نرسم شكلا لقيم الخطأ العشوائي على مدى الزمن، فإننا نتوقع الحصول على شكل يشبه الشكل رقم (٦-٣).



الشكل رقم (٦-٣)

وبالمقابل، قد تظهر قيم الخطأ العشوائي ارتباطا موجبا عبر الزمن نتيجة (على سبيل المثال) لبطء التكيف في السلوك الاقتصادي، هنا، قد نجد أن القيمة الموجبة لـ u تتبع بقيمة موجبة أخرى لـ u، وقيمة سالبة تتبعها قيم سالبة أخرى $\text{cov}(u_t, u_{t+1}) > 0$. وعلى افتراض أن قيم الخطأ العشوائي تتحدد بوساطة قيم خارجية (ستحدد بدقة أكثر فيما بعد) تتحول آثارها عشوائيا من الموجب إلى السالب، فإنه يمكننا أن نتوقع نمط قيم u عبر الزمن والذي يحتوي على نتائج مختلفة من قيم موجبة وسالبة. وعلى سبيل المثال إذا كانت القوى الخارجية تولد قيما موجبة لـ u، فإنها سوف تتبع بقيم موجبة أخرى، وإذا نتج عن هذه القوى قيم سالبة، فإنه سوف تتبعها قيم سالبة إضافية، وهكذا، يظهر مثل هذا النمط في الشكل رقم (٦-٤).



شكل (٦-٤)

وتعرف مشكلة الاعتماد المتبادل بين القيم المتتابعة للخطأ العشوائي بمشكلة الارتباط الذاتي autocorrelation. وسوف نبين أنه، في ظل افتراضات معينة، بالإضافة لافتراضاتنا السابقة، فإنه، عند أي قيمة معطاة للمشاهدة $E(u_t) = 0$، سوف لايكون الخطأ العشوائي مرتبطا بالمتغيرات المستقلة بحيث تظل مقدرات المعلمات غير متحيزة. وكما سنرى فيمابعد، فإن مشكلة الارتباط الذاتي تثير قلقا بشأن قيمة تباين المقدرات. وعلى نحو خاص، فإن الصيغ التي اشتقت للتباينات تصبح غير صحيحة إذا وجدت مشكلة الارتباط الذاتي. فإذا استمرينا في استخدام هذه الصيغ فسوف نصل إلى نسب t خاطئة مما يجعل اختباراتنا للفرضيات حول قيم المعلمات في نموذجنا خاطئة أيضا. ونتيجة لذلك فقد نقبل - مثلا - قيمة مقدرة لإحدى المعلمات على أساس أنها معنوية إحصائيا وهي، في الحقيقة، ليست مختلفة معنويا عن الصفر.

نموذج للانحدار الذاتي

لجعل مناقشتنا أكثر علمية دعنا نفترض أن عملية الانحدار الذاتي (أي الطريقة التي يرتبط بها خطأ عشوائي بآخر) تأخذ الشكل التالي:

$$u_t = \gamma u_{t-1} + \varepsilon_t, \qquad t = \cdots, -2, -1, 0, 1, 2, \cdots, \tag{6.18}$$

حيث إن ε_t هو متغير عشوائي موزع توزيعا طبيعيا وله قيمة متوسطة صفرية $E(\varepsilon_t) = 0$ ومستقل عن قيمته في أي فترة زمنية أخرى، لذا يكون $\text{cov}(\varepsilon_t, \varepsilon_s) = 0$ وله تباين ثابت $E(\varepsilon_t^2 = \sigma_t^2)$. ونفترض أيضا، (لأسباب سنعرضها فيما بعد) أن قيمة γ المطلقة أقل من الواحد الصحيح $|\gamma| < 1$. وباختصار، نفترض أن قيمة الخطأ العشوائي في أي فترة زمنية مرتبطة بالقيمة السابقة له مباشرة بوساطة نموذج انحدار خطي بسيط. لاحظ أنه يفترض أن المعادلة (6.18) تظل قائمة بالنسبة لجميع الفترات الزمنية الماضية والمستقبلية. وفي هذا النموذج، ترتبط قيم الخطأ العشوائي المتتابعة (أي u's) ببعضها البعض، وإن كان الارتباط ليس تاما وبالتحديد، إذا كانت γ موجبة فإن قيمة u_t سوف تكون مرتبطة إيجابيا مع قيمتها السابقة مباشرة، u_{t-1} أما

إذا كانت γ سالبة فإن هذا الارتباط يكون سالبـا، وتناظر الحالة الأخيرة مثالنا السابق عن الفرد الذي ينفق أكثر من المعتاد في إحدى الفترات، الفتـرة t مثـلاً، فسيحاول تعويض ذلك بانفاق أقل من المعتاد في الفترة اللاحـقـة. لاحـظ مـن المعادلة (6.18) أن الفرد، مع ذلك، قد لايقلل من الانفاق في الفترة التاليـة (t+1) لحدث آخر غير متوقع قد يجعله يزيد مرة أخرى انفاقه عن القـدر المعتـاد (أي ε_{t+1}).

سنبين أولا أنه، في ظل نموذج الانحدار الذاتي في المعـادلـة (6.18)، فـإن $E(u_t) = 0$ لاحظ من المعادلة (6.18) أن:

$$u_{t-1} = \gamma\, u_{t-2} + \varepsilon_{t-1}, \tag{6.19}$$

وبإحلال قيمة المتغيـر u_{t-1} من المعادلة (6.19) في المعادلة (6.18) نحصل علـى:

$$u_t = \gamma(\gamma u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \varepsilon_t + \gamma\varepsilon_{t-1} + \gamma^2 u_{t-2}. \tag{6.20}$$

ولما كانت المعادلة (6.18) تظل قائمة لجميع الفترات الزمنية، فبإحلال مايعـادل u_{t-2} وu_{t-3} وهلم جرا، فسنحصل على المقدار:

$$u_t = \varepsilon_t + \gamma\varepsilon_{t-1} + \gamma^2\varepsilon_{t-2} + \gamma^3\varepsilon_{t-3} + \cdots, \tag{6.21}$$

الذي لايتضمن قيما مبطـأة لـ u لأن معاملـه γ سيكون مرفوعا لقوى لانهائيـة (تساوي الصفر). وبأخذ القيمة المتوقعة لـ u_t في المعادلـة (6.21)، سيكون لدينا:

$$\begin{aligned} E(u_t) &= E(\varepsilon_t + \gamma\varepsilon_{t-1} + \gamma^2\varepsilon_{t-2} + \gamma^3\varepsilon_{t-3} + \cdots) \\ &= E(\varepsilon_t) + \gamma E(\varepsilon_{t-1}) + \gamma^2 E(\varepsilon_{t-2}) + \gamma^3 E(\varepsilon_{t-3}) + \cdots = 0. \end{aligned} \tag{6.22}$$

ولما كانت$E(\varepsilon_s) = 0$ حيث $s=t, t-1, t-2$ فإن القيمة المتوقعة للخطأ العشوائي لاتـزال تساوي الصفر. ويمكننا أن نرى، أنه بافتراض، كما يحدث عادة، استقلاليةε_t عن قيمة المتغير المستقـل، X_t مثلا، في جميع الفترات الزمنيـة (أي تكـونε_t مستقلة عنX_t لكلs,t)، فإنu_t ستكون مستقلة عن X_t، وهي هنا غير مرتبطة بـهـا. أي أنه، باستخـدام المـعـادلـة (6.21)، ينبغي أن تكون قـادرا عـلـى إثبـات أن: $\mathrm{cov}(u_t, X_t) = E(u_t, X_t) = 0$.

وعلى سبيل إشارة يمكن الرجوع إليها مستقبلا، نلاحظ أنه، طالما أن u_t توليفة خطية من ε_t ، ε_{t-1} ، ... وطالما أن جميع ε's مستقلة عن بعضها بعضا، فإن تباين u_t هو: *

$$\begin{aligned} \sigma_{u_t}^2 &= \sigma_\varepsilon^2 + \gamma^2 \sigma_\varepsilon^2 + (\lambda^2)^2 \sigma_\varepsilon^2 + (\gamma^3)^2 \sigma_\varepsilon^2 + \cdots \\ &= \sigma_\varepsilon^2 [1 + \gamma^2 + (\gamma^2)^2 + (\gamma^2)^3 + \cdots] \\ &= \frac{\sigma_\varepsilon^2}{1-\gamma^2} = \sigma_u^2, \end{aligned} \tag{6.23}$$

طالما أن جميع ε's لها التباين نفسه، وبافتراض أن $|\gamma| < 1$. من المعادلة (6.23)، نجد أن تباين u_t لايتضمن t، وكما هو الحال لـ ε's ، فإن جميع u_t's يكون لها التباين نفسه $\sigma_{u_t}^2 = \sigma_{u_s}^2 = \sigma_u^2$. ويمكننا أن نرى، أيضا، من المعادلة (6.23) لماذا نفترض أن $|\gamma| < 1$. فبدون هذا الافتراض، لاتتقارب السلسلة في المعادلة (6.23) ويصبح تباين u_t مالانهاية.

دعنا نفحص بعد ذلك التغاير للأخطاء العشوائية، فبالتعويض عن قيمة u_t من المعادلة (6.18) واستخدام المعادلة (6.23)، نحصل على:

$$\begin{aligned} E(u_t u_{t-1}) &= E[(\gamma u_{t-1} + \varepsilon_t) u_{t-1}] \\ &= \gamma E(u_{t-1)}^2) + E(\varepsilon_t u_{t-1}) \\ &= \gamma E(u_{t-1}^2) + 0 = \gamma \sigma_u^2, \end{aligned} \tag{6.24}$$

لأن u_{t-1} باستخدام المعادلة (6.21) يعتمد، فقط، على ε_{t-1} وقيمه المبطأة الأخرى ويعني هذا أن u_{t-1} و ε_t مستقلتان عن بعضهما بعضاً، ولذا فإن:

$$E(\varepsilon_t u_{t-1}) = \text{cov}(\varepsilon_t, u_{t-1}) = 0.$$

---

* استخدمنا هنا النظرية الأساسية لجمع المتوالية الهندسية التي تبين أنه إذا كان:

$$s = a(1 + \alpha + \alpha^2 + \alpha^3 + \cdots), \quad |\alpha| < 1$$

فأن

$$s = \frac{a}{1-\alpha}$$

فإذا كانت $\alpha \geq 1$ فإن المتوالية لن تتقارب.

وهكذا، نجد من المعادلة (6.24) أن شكل الانحدار الذاتي في معادلة (6.18) يخالف افتراضنا بأن التغاير بين الأخطاء العشوائية يساوي الصفر طالما أن γ لاتساوي الصفر.

تأثيره على تباينات المقدرات

سنبين الآن أن الارتباط الذاتي يؤدي إلى صيغ مختلفة لتباين المقدرات. افترض أن لدينا نموذجا من متغيرين من الشكل الموضح في الفصل الثاني:

$$Y_t = b_0 + b_1 X_t + u_t. \tag{6.25}$$

تذكر أنه وفقا لافتراضات النموذج (بما فيها التغاير الصفري بين الأخطاء العشوائية) وجدنا أن التباين الشرطي للمقدر $\hat{b}_1$ هو:

$$\mathrm{var}(\hat{b}_1) = \frac{\sigma_u^2}{\sum (X_t - \overline{X})^2}. \tag{6.26}$$

وقد استخدمنا، لاشتقاق هذه النتيجة، المعادلة (2.71) التي تشير إلى أن $\hat{b}_1$ يمكن التعبير عنه كمايلي:

$$\hat{b}_1 = b_1 + \frac{\sum (X_t - \overline{X}) u_t}{\sum (X_t - \overline{X})^2}. \tag{6.27}$$

عندئذ، دع $A = S(X_t - \overline{X})^2$ و $w_t = (X_t - \overline{X})$، ومن ثم، نعيد كتابة (6.27) في شكل مفصل على النحو:

$$\hat{b}_1 = b_1 + \frac{w_1}{A} u_1 + \cdots + \frac{w_n}{A} u_n. \tag{6.28}$$

وعلى افتراض أن قيم X's معطاة (ومن ثم، w's و A) فإن $\hat{b}_1$ تكون توليفة خطية من الأخطاء العشوائية. عند ذلك نشتق المعادلة (6.26) من المعادلة (6.28) باستخدام صيغة التباين لمجموع خطي من المتغيرات العشوائية غير المترابطة. ولكن، إذا كانت هناك مشكلة ارتباط ذاتي فليس بامكاننا استخدام هذه الصيغة لأن u_t وبالتالي، الحدود الموجودة في المعادلة (6.28) مترابطة. ويعني هذا أن المعادلة (6.26) لم تعد

الصيغة الصحيحة لتباين $\hat{b}_1$. ونتيجة لذلك فإن استخدام صيغنا التقليدية لاختبار الفرضيات لم تعد صحيحة كذلك.*

الوسط الحسابي للمقدرات

من السهل أن نرى أن الأرتباط الذاتي لايؤدي إلى تحيز في $\hat{b}_1$ حيث مايزال الخطأ العشوائي u_t، في ظل افتراضاتنا السابقة، مستقلاً عن جميع قيم X_t، ومن ثم، غير مرتبط بها (ومن ثم، w_t) وماتزال قيمته المتوقعة هي الصفر. فإذا أخذنا القيمة المتوقعة لـ (6.28) لأي قيم معطاة لـ X_t فإننا نحصل على:

$$E(\hat{b}_1)=b_1+\frac{w_1}{A}E(u_1)+\cdots+\frac{w_n}{A}E(u_n)=b_1. \tag{6.29}$$

وبالمثل، فإنه، باستخدام القاعدة $\hat{b}_0=\overline{Y}-\hat{b}_1\overline{X}$ ينبغي أن تكون قادرا على إثبات:

$$\begin{aligned} E(\hat{b}_0)&=b_0+b_1\overline{X}+E(\overline{u})-\overline{X}E(\hat{b}_1) \\ &=b_0 \end{aligned} \tag{6.30}$$

هنا ينبغي أن تكون السمة العامة لمشكلة الارتباط المتسلسل الذاتي واضحة، فعندما توجد هذه المشكلة فسيكون لدينا تغير منتظم في قيم الخطأ العشوائي للمشاهدات المتتابعة. مثل هذا النمط من التغير لايؤدي إلى إيجاد مقدرات متحيزة للمعلمات. إلا أن صيغنا للتباين لم تعد صحيحة، ونتيجة لذلك، وبدون نتائج إضافية، فإنه لايمكننا اختبار الفرضيات وإنشاء فترات ثقة. ومن الواضح أننا نحتاج إلى خصائص مرغوب فيها في منهجنا للتقدير. وأكثر من ذلك، فمن البديهي، طالما أن منهجنا في التقدير لايأخذ في الحسبان صراحة الارتباط المتسلسل الذاتي فإنه قد لايعطينا أكثر المقدرات دقة للمعلمات. أي أنه، إذا كان هناك نمط محدد للتغير بين الأخطاء

---

* باستخدام المعادلة (6.28)، ينبغي أن تكون لديك القدرة على إثبات أنه، في حالة الارتباط الذاتي، فإن تباين b (أي $E\left(\hat{b}_1-b_1\right)^2$) يتضمن القيمة المتوقعة لحدود ضرب التقاطع في المعادلة (6.28). وفي حالة غياب الارتباط الذاتي، تكون جميع حدود التقاطع هذه صفرية وبذلك، تسقط، تاركة لنا الصيغة (6.26) لتباين $\hat{b}_i$.

العشوائية، فينبغي أن نكون قادرين على الحصول على تقدير وتنبؤ افضل بأخذ هذه المعلومات الإضافية في حساباتنا. وسوف نفحص الآن كيف يمكن أن يحدث ذلك.

طريقة تقدير معممة

افترض أن نموذجنا يتكون من:

$$Y_t = b_0 + b_1 X_t + u_t \tag{6.31}$$

$$u_t = \gamma u_{t-1} + \varepsilon_t, \tag{6.32}$$

حيث $|\gamma| < 1$ و ε_t تحقق الافتراضات المعتادة السابق تكوينها كافة. مشكلتنا هنا تنحصر في كيفية استخدام المعلومات المعطاة في المعادلة (6.32) لتحسين مقدراتنا لمعلمات المعادلة (6.31).

افترض مبدئيا أن قيمة γ معلومة. فإذا أخذنا الصيغة المبطأة بالمعادلة (6.31) وضربناها بوساطة γ نحصل على:

$$\gamma Y_{t-1} = \gamma b_0 + \gamma b_1 X_{t-1} + \gamma u_{t-1}\,. \tag{6.33}$$

وبطرح المعادلة (6.33) من المعادلة (6.31) نحصل على:

$$Y_t - \gamma Y_{t-1} = (b_0 - \gamma b_0) + (b_1 X_t - \gamma b_1 X_{t-1}) + (u_t - \gamma u_{t-1}). \tag{6.34}$$

وبإعادة ترتيب الحدود في المعادلة (6.32)، نجد أن:

$$\varepsilon_t = (u_t - \gamma\, u_{t-1}),$$

التي، عند التعويض عن الحد الأخير في المعادلة (6.34)، نحصل على:

$$Y_t - \gamma Y_{t-1} = (b_0 - \gamma b_0) + (b_1 X_t - \gamma b_1 X_{t-1}) + \varepsilon_t\,.. \tag{6.35}$$

ويمكننا إعادة كتابة المعادلة (6.35) نتحصل على:

$$\begin{aligned} Y_t' &= B + b_1 X_t' + \varepsilon_t, \\ Y_t' &= Y_t - \gamma Y_{t-1}, \\ B &= b_0 - \gamma b_0, \\ X_t' &= X_t - \gamma X_{t-1}\,. \end{aligned} \tag{6.36}$$

لاحظ أننا، عند ترجمة المعادلة (6.31) إلى (6.36)، خسرنا مشاهدة واحدة بسبب الإبطاء والطرح في المعادلة (6.34).

وهكذا، تصبح المعادلة (6.31) الشكل المعتاد لنموذج الانحدار، خصوصا أن ε_t (وليس u_t) يحقق الافتراضات المرتبطة بخصائص الخطأ العشوائي كافة. لذلك، يمكننا أن نتبع منهج التقدير الذي وضعناه من قبل: نضع الشروط $\sum_{t=2}^{n}\hat{\varepsilon}_t = 0$ و $\sum_{t=2}^{n}(X_t' \ \hat{\varepsilon}_t) = 0$ وبتلك الوسيلة، يمكننا اشتقاق معادلتين طبيعيتين، وبحلهما، نحصل على مقدراتنا لـ $\hat{\beta}$ و $\hat{b}_1$، للمعلمتين في المعادلة (6.36). ويمكننا بعدئذ أن نقدر b_0 بوساطة:

$$\hat{b}_0 = \frac{\hat{B}}{1-\gamma}. \tag{6.37}$$

ويمكن، أيضا، إثبات أن:

$$\mathrm{var}(\hat{b}_0) = \left(\frac{1}{1-\gamma}\right)^2 \ \mathrm{var}(\hat{B}), \tag{6.38}$$

طالما أن $\hat{b}_0$ مرتبطة ارتباطا خطيا تاما مع $\hat{\beta}$. وطالما أن ε_t يحقق افتراضاتنا المعتادة كافة، فإنه يمكن الوصول إلى تباينات $\hat{\beta}$ و $\hat{b}_1$ بوساطة صيغنا المعتادة:

$$\mathrm{var}(\hat{B}) = \frac{\sigma_t^2 \sum_{t=2}^{n}(X_t')^2}{n' \sum_{t=2}^{n}(X_t' - \bar{X}')^2}, \qquad \mathrm{var}(\hat{b}_1) = \frac{\sigma_t^2}{\sum_{t=2}^{n}(X_t' - \bar{X}')^2} \tag{6.38}$$

حيث إن $n' = n-1$، نظرا لفقدان إحدى المشاهدات في عملية الإبطاء والطرح للحصول على X_t'.

يمكننا الآن - من حيث المبدأ، في الأقل - اختبار الفرضيات وانشاء فترات ثقة صحيحة إذا ماقدرنا معلماتنا الأساسية بوساطة المعادلة (6.36). ونشير، أيضا،

إلى أن المقدرات التي حصل عليها من المعادلة (6.36) كفاء، ويعني ذلك بدهيا*، أنه إذا كان حجم العينة كبيرا فإن تباينات مقدرات (6.36) ستكون أقل من تباينات أي مقدرات أخرى غير متحيزة لـ B و b_1 أو تساويها. والسبب الذي يمنعنا من القول أن مقدراتنا سوف يكون لها أقل التباينات بغض النظر عن حجم العينـة هـو أن طريقتنا في التقدير تتضمن، فقط، أن واحدة من المشاهدات. وأساساً، تتضاءل أهمية هذه المشاهدة بتزايد حجم العينة.**

وهكذا، فقد وجدنا طريقة للتقدير بصفات مرغوب فيها لدمج المعلومـات المتاحة حول العلاقة بين الأخطاء العشوائية ذاتها في منهجنا للتقدير. وبالتحديد، فقد حولنا نموذج الانحدار الذي يعاني من الارتباط الذاتي إلى نموذج يحقق كافة الافتراضات المرتبطة بنموذج الانحدار الأساسي (ومن بينها التغاير الصفـري بـين الأخطاء العشوائية) وبعد ذلك نطبق طرقنا المعتادة في التقدير. والصعوبة الـتـي نواجهها في تطبيق هذه الطريقة هي أن γ (وعمومًا) غير معلومة، ولذا ينبغي علينا أولا أن نقدر قيمة γ.***

ويمكننا أداء ذلك من خلال رؤية البواقي من معادلتنا الأصلية (6.31). ويتذكر أن $E(u_t) = 0$ وأن $\text{cov}(u_t, X_t) = 0$ يمكننا اشتقاق مقدرات غير متحيزة للمعلمات في

---

* ليس هذا تعريفا اصطلاحيا للمقدر الكفء، فمثل هذا التعريف خارج عن نطاق هذا الكتاب وغير ضروري لفهم النتائج التي سترد فيها بعد.

** تعرف الكفاءة، عادة، بدلالة مايسمى بمتوسط مربع الخطأ mean square error (م م خ). أي افترض أن $\hat{\alpha}$ هي مقدر لـ α. حينئذ، فإن م م خ لـ $\hat{\alpha}$ يساوي مجموع تباينها $\sigma^2_{\hat{\alpha}}$ ومربع تحيزها $E(\hat{\alpha}-\alpha)^2$. وفي حالتنا المذكورة أعلاه، طالما أن مقدراتنا غير متحيزة فإن م م خ يساوي التباين فقط. وعلى أي الأحوال، وبتجاهل قليل من الأشياء غير المهمة، إذا كانت $\hat{\alpha}$ مقدرا كفئا لـ α فإننا نتوقع (في حالة العينات الكبيرة) أن م م خ لـ سوف يكون أقل من م م خ لأي مقدر متسق لـ α أو يساويه.

*** نلاحظ في معادلات الانحدار أنه جميع متغيراتنا تأخذ شكل الفروق من الدرجة الأولى First difference form، أي بالنسبة لحالة المتغيرين الاثنين فقط $Y_t - Y_{t-1}$ يتم انحدارها على $X_t - X_{t-1}$. ومن المعادلة (6.36)، يمكن أن يتبين لنا أن مثل هذه المعادلة هي حالة خاصة لنموذج الانحدار الذاتي حيث تكون γ مساوية الواحد. لاحظ من المعادلة (6.23)، أنه بالنسبة لهذا النموذج، فإن تباين u سيكون لانهائيا. وأكثر من ذلك يكون هذا المنهج مقيدا جدا، طالما أن نتائجه مبنية على تحقق الشرط $\gamma=1$.

المعادلة (6.31) عن طريق فرض الشروط $\sum_{t=1}^{n}\hat{u}_t = 0$ و $\sum_{t=1}^{n}(\hat{u}_t X_t) = 0$ ويعطينا هذه المعادلات الطبيعية المعتادة.

$$\begin{aligned} \sum Y_t &= n\hat{b}_0 + \hat{b}_1 \sum X_t, \\ \sum (X_t Y_t) &= \hat{b}_0 \sum X_t + \hat{b}_1 \sum X_t^2, \end{aligned} \tag{6.40}$$

والتي يمكننا حلها للحصول على المقدرات غير المتحيزة $\hat{b}_0$ و $\hat{b}_1$. ويمكننا، حينئذ، أن نستخدم $\hat{b}_0$ و $\hat{b}_1$ للحصول على مقدر $(\hat{u}_t)$ لقيمة الخطأ العشوائي:

$$\hat{u}_t = Y_t - \hat{Y}_t = Y_t - (\hat{b}_0 + \hat{b}_1 X_t). \tag{6.41}$$

ولتقدير γ نعوض ببساطة عن قيمة $(\hat{u}_t)$ من المعادلة (6.41) في العلاقة المقترحة في المعادلة (6.32)، أي:

$$\hat{u}_t = \gamma \hat{u}_{t-1} + \varepsilon_t. \tag{6.42}$$

اعتبر المعادلة (6.42) نموذجا للانحدار. وبما أن ε_t مستقلة عن u_t دعنا نجعل ε_t غير مرتبطة بـ u_{t-1}. * وبالاستعانة بهذا الفرض يمكننا أن نقدر γ في المعادلة (6.42) بوساطة منهجنا العادي. وبالتحديد (على سبيل المراجعة) يمكننا إعادة كتابة المعادلة (6.42) على النحو التالي:

$$\hat{u}_t = \hat{\gamma} \hat{u}_{t-1} + \hat{\varepsilon}_t, \tag{6.43}$$

حيث إن $\hat{\varepsilon}_t(\hat{u}_t - \hat{\gamma}\hat{u}_{t-1})$ هو مقدر الخطأ العشوائي و $\hat{\gamma}$ هو مقدرنا لـ γ. ويعني افتراضنا $\text{cov}(\varepsilon_t \ \ \hat{u}_{t-1}) = 0$ أننا نفرض الشرط التالي (تذكر أن إحدى المشاهدات تفقد

---

* يتضائل الترابط بين ε_t و u_{t-1} ويؤول إلى الصفر، كلما ازداد حجم العينة إلى مالانهاية. أي أنه كلما كان حجم العينة صغيراً كان هنالك ترابط بين ε_t و $\hat{u}_{t-1}$ وذلك بسبب أن $\hat{u}_{t-1}$ تعتمد على كل من $\hat{b}_0$ و $\hat{b}_1$ اللذين يعتمدان، بدورهما، على جميع ε_s' مشتملة على ε_t. ولكن، مع زيادة حجم العينة، وطالما أن $\hat{b}_0$ و $\hat{b}_1$ متسقتان، فإنهما سوف تتقاربان في الاحتمال إلى b_0 و b_1 لذا، فإن $\hat{u}_{t-1}$ سوف تنحرف في النهاية عن u_{t-1} باحتمال يساوي الصفر. ويمكن افتراض صحة المعادلة (6.42)، باحتمال قدره الواحد الصحيح، فقط، في حالة ما إذا كان حجم العينة لانهائيا. باختصار، ينبغي علينا أن نعد (6.42) معادلة تقريبية (أوللعينات الكبيرة).

$\sum_{t=2}^{n}(\hat{\varepsilon}_t \ \hat{u}_{t-1})=0$. وللقيام بذلك نضرب حدود المعادلة (6.43) بوساطة $\hat{u}_{t-1}$ ، ثم نجمع على مدى العينة. وأخيرا نطبق شرطنا للحصول على:

$$\sum_{t=2}^{n}(\hat{u}_t u_{t-1})=\hat{\gamma}\sum_{t=2}^{n}(\hat{u}_{t-1})^2+\sum_{t=2}^{n}(\hat{\varepsilon}_t\hat{u}_{t-1})=\hat{\gamma}\sum_{t=2}^{n}(\hat{u}_{t-1})^2. \quad (6.44)$$

ومن المعادلة (6.44) يكون مقدرنا لـ γ هو:

$$\hat{\gamma}=\frac{\sum_{t=2}^{n}(\hat{u}_t\hat{u}_{t-1})}{\sum_{t=2}^{n}(\hat{u}_{t-1})^2} \quad (6.45)$$

يمكننا الآن أن نستخدم المنهج الذي وضعناه من قبل (باستخدام $\hat{\gamma}$ بدلا من γ) للحصول على مقدرات لنموذجنا للانحدار. في حالة نموذج من متغيرين، فسيكون لدينا:

$$\hat{b}_1=\frac{\sum_{t=2}^{n}(X_t^*-\overline{X})Y_t^*}{\sum_{t=2}^{n}(X_t^*-\overline{X}^*)^2} \quad (6.46)$$

وأيضا

$$\hat{b}_0=\frac{\hat{B}}{1-\hat{\gamma}},$$

وحيث

$$X_t^*=X_t-\hat{\gamma}X_{t-1}, \quad (6.47)$$
$$Y_t^*=Y_t-\hat{\gamma}Y_{t-1},$$
$$\hat{B}=\overline{Y}^*-\hat{b}_1\overline{X}^*.$$

وبالمثل ووفقا لمناقشتنا اللاحقة ستكون صيغتا التباين هي:

$$\operatorname{var}(\hat{b}_0)=\frac{1}{(1-\hat{\gamma})^2}\operatorname{var}(\hat{B}),$$

$$\operatorname{var}(\hat{b}_1)=\frac{\sigma_t^2}{\sum_{t=2}^{n}(X_t^*-X^*)^2} \tag{6.48}$$

حيث

$$\operatorname{var}(\hat{B})=\frac{\sigma_\varepsilon^2\sum_{t=2}^{n}(X_t^*)^2}{n'\sum_{t=2}^{n}(X_t^*-\overline{X}^*)^2} \tag{6.49}$$

حيث $n'=n-1$. وباختصار طالما أننا حولنا معادلتنا باستخدام $\widehat{\gamma}$، فإننا سوف نعامل $\widehat{\gamma}$ كما لو أنها γ وسنستخدم جميع صيغنا المعتادة. وهذا يتضمن تقديـر σ_ε^2 في صيغ التباين المذكورة أعلاه. وبالتحديد ينبغي أن يكون واضحا من (6.36) أن مقدرنا σ_ε^2 سيكون:

$$\hat{\sigma}_\varepsilon^2=\sum_{t=2}^{N}\frac{(Y_t^*-\hat{B}-\hat{b}_1X_t^*)^2}{(n'-2)} \tag{6.49}$$

لنبحث الآن قضايا اختبار الفرضيات وتكوين فترات الثقة. من (6.45) نجد أن $\widehat{\gamma}$ تعتمد على الأخطاء العشوائية المقدرة. ومن (6.46) نجد أن $\hat{b}_1$ تعتمد اعتمادا غير خطي على $\widehat{\gamma}$ وينتج عن ذلك أن $\hat{b}_1$ تعتمد - من ضمان اشياء أخرى - على الأخطاء العشوائية المقدرة، إضافة إلى ذلك وطالما أن الأخطاء العشوائيـة [انـظـر المعادلة و (6.25)]، فإن $\hat{b}_1$ تعتمد اعتمادا غير خطي على الأخطاء العشوائية ونتيجة لذلك، فإن $\hat{b}_1$ ليست موزعة توزيعا طبيعـاً، كما أن النسبـة $\left(\hat{b}_1-b_1\right)/\hat{\sigma}_{b_1}$ ليست متغير t بدرجات حريـة n-2، حيث إن $\hat{\sigma}_{b_1}$ هو الانحراف المعياري المقدر لـــ $\hat{b}_1$. ويمكن الوصول إلى نتائج مشابهة لكل من $\hat{B}$ و $\hat{b}_0$ طالما أنهما يعتمدان اعتمادا غيـر خطي أيضا، على $\hat{\gamma}$.

ولحسن الحظ، يمكن اختبار الفرضيات بطريقة تقريبية أو تكون فترات ثقة تقريبية عن طريق افتراض أن النسب $\left(\hat{b}_0 - b_0\right)/\hat{\sigma}_{\hat{b}_0}$ و $\left(\hat{b}_1 - b_1\right)/\hat{\sigma}_{\hat{b}_1}$ وأخير $\left(\hat{B} - B\right)/\hat{\sigma}_B$ هي متغيرات طبيعية معيارية. فإذا كان حجم العينة لانهائياً، فإن هذه النسب، في الحقيقة هي متغيرات طبيعية معيارية. وفي حالة العينة النهائية (المحدودة) يمكن أن ينظر إلى افتراض الطبيعية على أنه افتراض تقريبي. ولذلك، فإن اختبار الفرضيات وتكوين فترات الثقة المبنية على افتراض الطبيعية ينبغي أن ينظر إليها على أنها تقريبية. ونلاحظ (بالمثل) أن صيغ التباين السابقة (عند استخدام $\hat{\lambda}$ بدلا من λ) هي تقريبية، أيضا، بمعنى أنها تكون صحيحة فقط في حال، العينة اللانهائية.

ينتج عن المناقشة السابقة أن فترة ثقة تقريبية 95% لـ$\hat{b}_1$ ستكون$\left(\hat{b}_1 \pm 1.96\ \sigma_{\hat{b}_1}\right)$ فإذا كنا نرغب في اختبار الفرضية $H_0 : b_1 = 0$ مقابل الفرضية البديلة $H_1 : b_1 \neq 0$ عند مستوى معنوية 5%. فإننا سنقبل فرضية العدم H_0* إذا كان $\left|\hat{b}_1 / \hat{\sigma}_{\hat{b}_1}\right| < 1.96$ ونرفضها إذا لم تتحقق هذه المتباينة. أما إذا كان الاهتمام بما نستمر في تسميته نسبة $t\left(\hat{b}_1 / \hat{\sigma}_{\hat{b}_1}\right)$، فإن الاختلاف الوحيد هو أن القيمة الحرجة الدقيقة ترتبط بالتوزيع الطبيعي بدلا من توزيع t. ولهذا السبب نجد الباحثين يحسبون ويكتبون نتائج الانحدار غالبا في شكل نسب تسهيلا لقرّائهم.

ونشير هنا إلى أن مقدراتنا المبنية على $\hat{\gamma}$ لم تعد غير متحيزة، ولكنها تتسم بصفة مرغوب فيها وهي الاتساق**، إضافة إلى ذلك، فإن هذه المقدرات تتسم بالكفاءة، لذلك لاتوجد مقدرات متسقة أخرى أفضل لمعلمات النموذج (في الأقل في العينات الكبيرة).

* لانستطيع (كما يذكر المؤلف) القول بقبول فرضية العدم (اصطلاحيًا وإنما القبول، فقط، يكون للفرضية البديلة، ويمكننا، في حالة عدم تحقق الفرضية البديلة، أن نذكر أننا لم نستطيع رفض فرضية العدم، وهناك فرق واضح بين عدم القدرة على رفض فرضية العدم وبين قبولها. ملحوظة المترجم.

** لمعرفة المزيد من القضايا المتضمنة في التحليل السابق يرجى الرجوع إلى:
Arthur S. Goldberger, *Econometric Theory* (New York: Wiley, 1964) Chap. 6.

حال نموذج الانحدار المتعدد

يمكن، بسهولة، توسيع نطاق الطريقة السابقة ليشتمل على معالجة مشكلة الارتباط الذاتي في حالة الانحدار المتعدد. افترض (مثلاً) أننا قد احتفظنا بالافتراضات السابقة كافة ماعدا أن لدينا الآن متغيرات مستقلة عددها k، حينئذ يأخذ نموذجنا الشكل:

$$\begin{aligned} Y_t &= b_0 + b_1 X_{1t} + \cdots + b_k X_{kt} + u_t, \\ u_t &= \gamma u_{t-1} + \varepsilon_t. \end{aligned} \tag{6.50}$$

سنقدر أولا معلمات النموذج (b_0، b_1، ...، b_k) بطريقتنا المعتادة ثم نقدر بعد ذلك الخطأ العشوائي بوساطة $\hat{u}_t = Y_t - \hat{Y}_t$. ثم نقدر γ بعد ذلك بوساطة (6.45)، ونحول بعدها المتغير التابع إلى Y_t^*، $\left(Y_t^* = Y_t - \hat{\gamma} Y_{t-1}\right)$، وكذلك المتغيرات المستقلة إلى $X_t^* = X_{it} - \hat{\gamma}\ X_{i(t-1)}$ عندها يكون لدينا نموذج الانحدار التالي:

$$Y_t^* = B + b_1 X_{1_t}^* + \cdots + b_k X_{k_t}^* + \varepsilon_t , \tag{6.51}$$

وسوف نقدر B، b_1، ...، b_k وأيضا، تباينات مقدراتنا بطرق التقدير المعروفة، ومرة أخرى، فإن مقدرات المعلمات ستكون متحيزة إلا أنها متسقة. وأخيرا سنختبر الفرضيات أو فترات الثقة عن طريق افتراض أن النسب $\left(\hat{b}_i - b_i\right) / \hat{\sigma}_{\hat{b}_i}$ هي متغيرات طبيعية معيارية، وكما عرفنا في الحالة السابقة، فإن اختبارات الفرضيات هذه وتكوين فترات الثقة أو حتى صيغ التباين لمقدرات المعلمات تكون صحيحة تماما في حالة العينة اللانهائية. ولذا، ينبغي أن نفسر نتائجنا بأنها تقريب للعينات النهائية.

ونشير في هذه العجالة إلى أن هذه الطريقة التي شرحناها ليست هي الطريقة الوحيدة لتصحيح الارتباط الذاتي. فهناك طريقتان تستخدمان استخداما واسعا وهما طريقتا كوكرين أوركت Cochrane-Orcutt وهيلدروث لو Hildreth-Lu. وبالنسبة للنموذج الحالي، فهذه الطرق تعادل الطريقة التي اتبعناها من حيث إن هذه الطرق تنتج مقدرات تكون خواص العينات الكبيرة منها متماثلة مع مقدراتنا، أي أن المقدرات تتسم بالاتساق والكفاءة.

اختبار ديربن – واتسون للارتباط الذاتي

دعنا نفترض أنه إذا كانت الأخطاء العشوائية في نموذج الانحدار مرتبطة ذاتياً، فإن علاقة الارتباط هذه تأخذ الشكل الموجود في النموذج (6.32). ولدينا الآن طريقة لمعالجة مشكلة الارتباط الذاتي. ولكننا لم نوجد الوسائل التي نتعرف بها على ما إذا كان لدينا مشكلة ارتباط ذاتي أم لا. بدلا من ذلك كان منهجنا السابق يعتمد على افتراض أنه في (6.32) تكون $\gamma \neq 0$، ومن الواضح أن من الأفضل أن نختبر هذه الفرضية.

وإحدى الطرق المباشرة لأداء ذلك هو أخذ $\gamma = 0$ على أنها فرضية العدم ثم فحص إمكانية رفض هذه الفرضية في صالح الفرضية البديلة $\gamma \neq 0$ عند مستوى من المعنوية. وننشئ فترة للثقة (بما يماثل طريقتنا في الفصل الثالث) لمقدرنا لـ γ. فإذا تضمنت فترتنا للثقة لـ $\hat{\gamma}$ الصفر، فسوف نقبل فرضية العدم بأن $\gamma = 0$، وإذا لم تتضمنه (بسبب أن $\hat{\gamma}$ موجبة أو سالبة بدرجة كافية) فسنرفض فرضية أن $\gamma = 0$، وفي هذه الحالة الأخيرة، نكون قد قبلنا فرضية أن الخطأ العشوائي مرتبط ذاتيا.*

يوجد لحسن الحظ اختبار لهذا النوع من الارتباط الذاتي طور ج. ديربن– و ج واتسون**، ويستخدم الاختبار مايشار إليه عادة باحصائية d لديربن واتسون، والمبنية على مجموع مربع الفروق في القيم المتتابعة للأخطاء العشوائية المقدرة.

$$d = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2}. \qquad (6.52)$$

* لمناقشة هذه الطرق أنظر:

S. Goldfeld and R. Quandt. *Non-Linear Methods in Econometrics* (Amsterdam: North Holland, 1972), pp. 183-186.

** يرجع إلى:

J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least-Squares Regression", parts I and II, *Biometrika 37* (1950), pp. 409-428 and 38 (1951), pp. 159-178. See also the discussion of the test in Arthur S. Goldberger, *Econometric Theory* (New York: Wiley, 1964), pp. 243-244; and in J. Johnston. *Econometric Methods*, 2nd ed. (New York: McGraw-Hill, 1972), pp. 249-254.

وبديهيا يمكننا أن نرى أنه إذا كان لدينا ارتباط ذاتيٌ موجب فإن القيم المتتابعة للأخطاء العشوائية سوف تميل للاقتراب من بعضها بعضا اقترابا غير عادي، فقيمة موجبة للخطأ العشوائي في الفترة t سوف تتبعها، على الأرجح، قيمة موجبة أخرى في الفترة $t+1$. ويعني هذا أن الحدود في بسط (6.52) سوف تكون صغيرة نسبيا ولذلك نتوقع أن الارتباط الذاتي الموجب ينتج عنه قيم صغيرة لـ d. وعلى العكس فإن الارتباط الذاتي السالب سيؤدي إلى ايجاد اختلافات كبيرة بين القيم المتتابعة لـ u_t. وتكون علامة هذا النوع من الارتباط الذاتي هي قيمة كبيرة غير عادية لـ d.

دعنا نفترض صحة افتراضنا في نموذج الانحدار الأصلي بأن $E(u_s\, u_t) = 0$، حيث $s \neq t$، لذا لاتوجد مشكلة الارتباط الذاتي. في هذه الحال، نتوقع أن التباين بين البواقي المقدرة $\hat{u}_t$ و $\hat{u}_{t-1}$ يكون صفرا بالتقريب. عندما يكون ذلك صحيحا فإنه يتبين لنا عن طريق فك بسط احصائية d في (6.52) أنه إذا كانت n كبيرة، فإن d ينبغي أن تكون قريبة من 2.*

$$d = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2} \doteq \frac{2\sum_{t=2}^{n} \hat{u}_t^2 - 2\sum_{t=2}^{n} (\hat{u}_t \hat{u}_{t-1})}{\sum_{t=1}^{n} \hat{u}_t^2}$$

$$= \frac{\left[2\sum_{t=2}^{n} \hat{u}_t^2 / (n-1)\right] - \left[2\sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1} / (n-1)\right]}{\left[\sum_{t=1}^{n} \hat{u}_t^2 / (n-1)\right]} \tag{6.53}$$

طالما أن $\Sigma_{t=2}^{n}(\hat{u}_t\, \hat{u}_{t-1})/n-1 = 0$، $\Sigma_{t=2}^{n} u_t^2 = \Sigma_{t=1}^{n} \hat{u}_t^2$ و $\Sigma_{t=2}^{n} u_t^2 = \Sigma_{t=2}^{n} \hat{u}_{t-1}^2$ وحقيقة فإنه (وعلى سبيل التعميم إلى حد ما) ينتج من (6.53) أنه إذا كانت n كبيرة. وإذا

* إذا كان حجم العينة n لانهائيا، فإن قيمة d ستصبح 2 باحتمال قدره الواحد الصحيح.

افترضنا أن نموذجنا المرتبط ذاتيا السابق، $u_t = \gamma u_{t-1} + \varepsilon_t$:

$$\begin{aligned} d &\doteq \frac{2\operatorname{var}(u_t) - 2\operatorname{cov}(u_t, u_{t-1})}{\operatorname{var}(u_t)} \\ &= \frac{2\sigma_u^2 - 2\gamma\sigma_u^2}{\sigma_u^2} = 2(1-\gamma), \end{aligned} \tag{6.54}$$

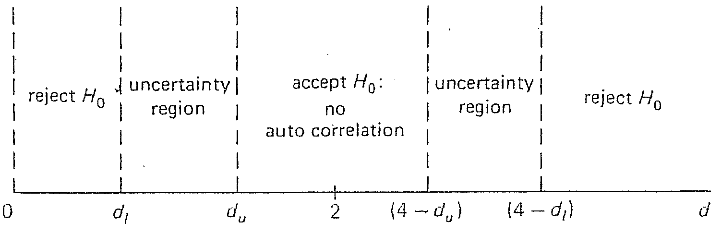
وطالما أن $\hat{u}_t$ مقدر لـ u_t، ومن (6.42)، $\operatorname{cov}(u_t u_{t-1}) = \gamma\sigma_2^2$ وباختصار فإننا نرى أن:

$$\begin{aligned} &\gamma = 0 \quad \text{توحي بأن} \quad d \doteq 2 \\ &\gamma = 1 \quad \text{توحي بأن} \quad d \doteq 0, \\ &\gamma = -1 \quad \text{توحي بأن} \quad d \doteq 4. \end{aligned} \tag{6.55}$$

يوحي لنا كل ماسبق أننا إذا أردنا اختبار فرضية العدم وهي عدم وجود ارتباط ذاتي $H_0 : \gamma = 0$ إزاء الفرضية البديلة بوجوده $H_1 : \gamma \neq 0$ فإننا سنقبل H_0 إذا كانت قيمة d قريبة قربا كافيا من 2، وسوف نقبل الفرضية البديلة إذا لم تكن d كذلك. وبالمقابل فإن قيم d التي تكون قريبة من الصفر أو من 4 ستقودنا إلى قبول الفرضية البديلة $H_1 : \gamma \neq 0$.

ولسوء الحظ، وبسبب وجود خواص إحصائية معينة للإحصائية d، فإن المشكلة أكثر تعقيدا. وبخاصة أن المناطق المناظرة لقبول أو رفضها فرضية العدم (للإحصائية d) بعدم وجود ارتباط ذاتي لاتستنفذ جميع القيم الممكنة لـ d. لذلك يوجد مدى من القيم لايمكننا خلالها أن نقبل أو نرفض H_0. وبالتحديد، في حالة اختبار ديربن واتسون ذو الطرفين مع $H_0 : \gamma = 0$ ازاء $H_1 : \gamma \neq 0$، توجد لدينا مجموعة من خمس مناطق لقيم d كما تظهر في الشكل (٦-٥). فإذا كانت d أقل من d_L أو أكبر من $(4 - d_L)$ فإننا نرفض فرضية العدم في صالح الفرضية البديلة، ويتضمن هذا وجود الارتباط الذاتي. وعلى العكس، إذا كانت قيمة d قريبة من 2، أو بصورة أكثر دقة بين d_u و $(4 - d_u)$ نقبل فرضية العدم بعدم وجود ارتباط ذاتي. أما إذا كانت

قيمة d تنحصر بين d_L و d_u أو بين $(4 - d_u)$ أو $(4 - d_L)$ فإن اختبار ديربن واتسون يصبح غير حاسم لأنه عند هذه القيم من d لايمكننا عند مستوى محدد من المعنويـة، أن نستنتج وجود الارتباط الذاتي أو عدم وجود بين الأخطاء العشوائية. أي أنه، على العكس من اختباراتنا السابقة، يتضمن اختبار ديربن واتسون (وبسبب صعـوبـات إحصائية معينة) مناطق عدم تأكد.



الشكل رقم (٦-٥)

وتتبع طريقة اختبار الذيل الواحد مباشرة المنهج السابق. افترض مثلا أننـا مهتمون بالفرضية $H_0 : \gamma = 0$ مقابل الفرضية $H_1 : \gamma > 0$. حينئذ سوف نقبل H_0 إذا كانت d بعيدة «بعدا كافيـاً» عن الصفر، واصطلاحيـاً، وبدلالة الشكل (٦-٥) سوف نقبل H_0 إذا كانت $d > d_u$ ونرفضها إذا كانـت $d < d_L$، ويكون الاختبار غير حاسم إذا كانت $d_L < d < d_u$. وبالمثل إذا كانت الفرضية البديلة هـي $\gamma < 0$، فإننـا سنقبل H_0 إذا كانت $d < 4 - d_u$ ونرفضها إذا كانت $d > 4 - d_L$، ويكون الاختبار غير حاسم إذا كانت $4-d_u < d < 4-d_L$.

يوجد في نهاية هذا الكتاب جدول بقيـم d_u و d_L في الجدول الإحصائي ٤ ولإيجاد قيمة معينة لـ d_u و d_L للمشكلة الحالية، نحتاج لمعرفة مستوى المعنوية، وما إذا كان الاختبار ذا طرف واحد أم اثنين. وكذلك حجـم العينة، وأخيرا، عـدد

المتغيرات المستقلة (k’) في معادلة الانحـدار.* وبالرجوع إلى الجدول الإحصائي٤ نجد أنه، على سبيل المثال، وعند مستوى معنـويـة 5% ($\alpha = 0.05$) في اختبـار ذو الطرفين، فإنه إذا كان لدينا 50 مشاهدة (n=50)، ومعادلة لها ثلاثة متغيرات مستقلة (k’=3)، حيـنـئـذ، فـإن$d_L = 1.34$ و $d_u = 1.54$. وعند الحـصـول عـلـى هـذه الأرقـــام، نـــلاحـــظ أن قـــيـــم d_u وd_L المعطاة في الجدول والتي تناظر 0.025 مستوى معنوية (أي 0.025 في كل طرف). في هذه الحال، تكون المناطق الخمس في الشكل (٦-٥) هي:

$$
\begin{aligned}
&a.(0,d_l)=(0,1.34),\\
&b.(d_1,d_u)=(1.34,1.59),\\
&c.(d_u,4-d_u)=(1.59,2.41),\\
&d.(4-d_u,4-d_l)=(2.41,2.66),\\
&e..(4-d_l,4)=(2.66,4.00).
\end{aligned}
$$

بعد ذلك، نستخدم المعادلـة (6.52) لحساب القيمة الفعلية لـ d من القيم المشاهـدة لمتغيراتنا ونحدد في أي من المناطق الخمس تقع هذه القيمة لمعرفة ما إذا كان علينا أن نتخوف من وجود مشكلة الارتباط الذاتي.

تطبيق

قد يكون من المفيد لمراجعة معالجتنا للارتباط الذاتي أن ننهي مناقشتنا بمثال توضيحي يشتمل على أرقام واقعية. افترض أننا نرغب في تقدير دالة استـهـلاك بسيط تأخذ الشكل:

$$C_t = b_0 + b_1 Y_{dt} + u_t, \tag{6.56}$$

حيث إنC_t (كما في المثال السابق) الإنفاق الاستهلاكي وY_{dt} الدخل المتاح، ولدينا مجموعة من المشاهدات السنوية عن الاستهلاك الاجمالي، والدخل المتاح للولايات

---

* لاتشتمل k على الحد الثابت. ويمكن أن تعرف k’ (مثلاً) بأنها عدد معلمات الميل في الانحدار.

المتحدة الأمريكية للسنوات١٩٥١-١٩٦٩م، التي تظهر في الجدول رقم (٦-١). فإذا أجرينا انحدارا لـ C على Y_d باستخدام البيانات المتاحة في الجدول رقم (٦-١) نحصل على:

$$\hat{C}_t = 3.29 + 0.906 Y_{dt}, \qquad n = 19, \tag{6.57}$$
$$(1.5) \quad (162.0) \qquad R^2 = 0.999,$$

حيث تظهر نسب t أسفل تقديرات المعاملات. تفسر هذه المعادلة، بوضوح، معظم التغير في الاستهلاك (كما يظهر من R^2 التي تقترب من الواحد الصحيح)، وأكثر من ذلك، فإن تباين $\hat{b}_i$ صغير جداً، كما يستدل عليه من القيمة الهائلة لنسبة t المناظرة.

جدول رقم (٦-١) ببلايين الدولارات الأمريكية

السنة	الانفاق الاستهلاكي	الدخل المتاح
١٩٥١	٢٠٦,٣	٢٢٦,٦
١٩٥٢	٢١٦,٧	٢٣٨,٣
١٩٥٣	٢٣٠,٠	٢٥٢,٦
١٩٥٤	٢٣٦,٥	٢٥٧,٤
١٩٥٥	٢٥٤,٤	٢٧٥,٣
١٩٥٦	٢٦٦,٧	٢٩٣,٢
١٩٥٧	٢٨١,٤	٣٠٨,٥
١٩٥٨	٢٩٠,١	٣١٨,٨
١٩٥٩	٣١١,٢	٣٣٧,٣
١٩٦٠	٣٢٥,٢	٣٥٠,٠
١٩٦١	٣٣٥,٢	٣٦٤,٤
١٩٦٢	٣٥٥,١	٣٨٥,٥
١٩٦٣	٣٧٥,٠	٤٠٤,٦
١٩٦٤	٤٠١,٢	٤٣٨,١
١٩٦٥	٤٣٢,٨	٤٧٣,٢
١٩٦٦	٤٦٦,٣	٥١١,٩
١٩٦٧	٤٩٢,١	٥٤٦,٣
١٩٦٨	٥٣٦,٢	٥٩١,٠
١٩٦٩	٥٧٩,٦	٦٣٤,٢

المصدر: التقرير الاقتصادي الرئيسي، واشنطن، يناير ١٩٧٢، صفحة ٢١٢.

دعنا الآن نفحص القيم المقدرة للأخطاء العشوائية لنرى ما إذا كانت هنـاك أي إشارة لوجود الارتباط الذاتي. وتحسب معظم برامج الحاسوب لتحليل الانحدار قيمـة إحصائية d لديربان واتسون، ولذلك؛ فإن التساؤل يمكن الاجابة عليه مباشرة. ولكـن لنتعرف علـى هذا المنهج فإن الجدول رقم (٦-٢) يظهر التسلسل الفعلي للحسابـات. ولحساب الأخطاء العشوائية، نستخدم أولا المعادلة المقدرة (6.57) لحساب القيمة المتوقعة للاستهلاك كل سنة ثم نطرح هذه القيمة من الاستهلاك الفعلي للحصول على تقديرات الأخطاء العشوائية التي تظهر في العمود الرابع. فإذا نظرت إلى الأرقام في هذا العمود فسوف تلاحظ أن هناك تعاقبات للأخطاء العشوائية المقدرة، فقيم سالبة تتبعها سلسلة من القيم الموجبة للأخطاء وهذه تجعلنا نشك فورا لأن ذلك يوحي بوجود ارتباط متسلسـل موجب بين الأخطاء العشوائية وتجعلنا نتوقع قيمة منخفضة للإحصائية d. وفي الحقيقة فإن قيمة الإحصائية d منخفضة جدا عن ٢ حيث تعادل ١,٠١ فإذا مارجعنا مرة ثانية للجدول الإحصائي ٤ نجد أنه وحسب اختبار ديربن واتسون للإحصائية d (الاختبار ذو الطرفين*. أنه إذا كانت $\alpha = 0.5, k=1, n=19$ فإن قيمة الحد الأدنى، هي 1.06، وإحصائيا تقع d أسفل هذا الحد الأدنى. ولذلك نرفض فرضية العدم بعدم وجود ارتبـاط ذاتـي لصالح الفرضية البديلة بوجود الارتباط الذاتي بين الأخطاء العشوائية ($\gamma \neq 0$).

ولاستخدام المنهج الموضح سابقا لتصحيح النموذج من الارتباط الذاتي، ينبغي أن نقدر أولا العلاقة بين الأخطاء العشوائية. نفترض أن هذه العلاقة تأخذ الشـكـل التالي:

$$u_t = \gamma u_{t-1} + \varepsilon_t$$

حيث إن ε_t تحقق الافتراضات كافة التي وضعناها من قبل، وبأخذ القيم المقدرة للأخطاء العشوائية في الجدول رقم (٦-٢) نستخدم الصيغة الموجودة في (6.45) لتقدير قيمة γ:

* يجب تكوين الافتراضات قبل اختبار النتائج، ولهذا السبب نستخدم الاختبار ذو الطرفين، ويتضمن ذلك أنه قبل اختبار البواقي فعلياً، لايكون لدينا سبب مقنع لتوقع أنه إذا كانت لدينا مشكلة ارتباط متسلسل ذاتي فإن ذلك الارتباط سيكون موجباً.

جدول (٦-٢)

Year	Actual Consumption (C_t)	Predicted Consumption ($\hat{C}_t$)	$\hat{u}_t = C_t - \hat{C}_t$	$\hat{u}_t^2$	$\hat{u}_{t-1}$	$\hat{u}_t - \hat{u}_{t-1}$	$(\hat{u}_t - \hat{u}_{t-1})^2$
1951	206.3	208.6	−2.3	5.2			
1952	216.7	219.2	−2.5	6.2	−2.3	−0.2	0.0
1953	230.0	232.1	−2.1	4.6	−2.5	0.4	0.1
1954	236.5	236.5	0	0	−2.1	2.1	4.6
1955	254.4	252.7	1.7	2.9	0	1.7	2.8
1956	266.7	268.9	−2.2	5.0	1.7	−3.9	15.3
1957	281.4	282.8	−1.4	1.9	−2.2	0.8	0.7
1958	290.1	292.1	−2.0	4.1	−1.4	−0.6	0.4
1959	311.2	308.9	2.3	5.4	−2.0	4.3	18.8
1960	325.2	320.4	4.8	23.1	2.3	2.5	6.2
1961	335.2	333.4	1.8	3.1	4.8	−3.0	9.3
1962	355.1	352.4	2.7	7.4	1.8	1.0	0.9
1963	375.0	369.9	5.1	26.5	2.7	2.4	5.8
1964	401.2	400.2	1.0	1.0	5.1	−4.2	17.2
1965	432.8	432.0	0.8	0.6	1.0	−0.2	0.0
1966	466.3	467.1	−0.8	0.6	0.8	−1.6	2.4
1967	492.1	498.2	−6.1	36.7	−0.8	−5.4	28.8
1968	536.2	538.7	−2.5	6.4	−6.1	3.6	13.0
1969	579.6	577.9	1.7	3.0	−2.5	4.3	18.2
				$\sum \hat{u}_t^2 = 143.7$			$\sum (\hat{u}_t - \hat{u}_{t-1})^2 = 144.5$

$$d = \frac{\sum (\hat{u}_t - \hat{u}_{t-1})^2}{\sum \hat{u}_t^2} = \frac{144.5}{143.7} = 1.01$$

[a] Figures may not sum precisely due to rounding.

$$\hat{\gamma} = \frac{\sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^{n} \hat{u}_{t-1}^2} = 0.48. \qquad (6.58)$$

وباستخدام قيمة مقدرة لـ γ مساوية لـ 0.48، نحسب بعد ذلك:

$$C_t^* = C_t - \hat{\gamma} C_{t-1} = C_t - 0.48 C_{t-1},$$
$$Y_{dt}^* = T_{dt} - \hat{\gamma} Y_{d(t-1)} = Y_{dt} - 0.48 Y_{d(t-1)}.$$

وبتكوين انحدار لـ C^* على Y_d^* نجد أن:

$$\hat{C}_t^* = \underset{(1.0)}{2.12} + \underset{(98.9)}{0.905} Y_{dt}^*, \qquad n = 18, \quad R^2 = 0.998. \qquad (6.59)$$

وأخيرا فإن تقديراتنا للحد الثابت ولتباينه هي:

$$\hat{b}_0 = \frac{\hat{B}}{1-\hat{\gamma}} = \frac{2.12}{(1-0.48)} = 4.08$$

$$\hat{\text{var}}(\hat{b}_0) = \frac{1}{(1=\hat{\gamma})^2} \hat{\text{var}}(\hat{B}) = \frac{1}{(1-0.48)^2}(4.2) = 15.5.$$

وتكون معادلتنا المقدرة والمصححة من الارتباط الذاتي هي:*

$$\hat{C}_t = \underset{(1.0)}{4.08} + \underset{(98.9)}{0.905} Y_{dt}$$

لاحظ أنه، على الرغم من أن القيمة المقدرة للميل الحدي للاستهلاك (م ح س) b_1، في هذه الحال، تأخذ فعليا قيمة معادلة الانحدار العادية نفسها (6.50) فإن نسبة t المناظرة أصبحت أقل بدرجة كبيرة عندما صححنا مشكلة الارتباط الذاتي، وفي حالات أخرى، قد يحسم هذا الخلاف بين رفض فرضية العدم بوجود القيمة الصفرية للمعلمة أو قبولها.

* حسبت نسبة t للحد الثابت عن طريق قسمة الانحراف المعياري لـ $\hat{b}_0$ (الجذر التربيعي لتباينة) على $\hat{b}_0$.

الارتباط الذاتي والمتغيرات التابعة المبطأة

يعد اختبار ديربن واتسون غير صحيح، في حال، احتواء نموذج الانحدار على قيمة مبطأة للمتغير التابع باعتبارها واحدا من المتغيرات المستقلة. ولفهم المشاكل المتضمنة افترض النموذج التالي:

$$Y_t = b_0 + b_1 X_{1t} + \cdots + b_k X_{kt} + aY_{t-1} + u_t \tag{6.60}$$

$$u_t = \gamma u_{t-1} + \varepsilon_t, \qquad t = \cdots, -2, -1, 0, 1, 2, \cdots \tag{6.61}$$

حيث إن ε_t لها قيمة متوقعة صفرية $E(\varepsilon_t) = 0$ وتباين ثابت $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ و $E(\varepsilon_t\ \varepsilon_s) = 0$ إذا كانت $t \neq s$. نفترض، أيضا، أن $|\gamma| < 1$، وأن $|a| < 1$ يتماثل هذا النموذج مع النموذج الموجود في (6.50) باستثناء أنه في (6.60)، يحتوي على متغير مستقل Y_{t-1} بينما لايحتوي (6.50) على ذلك المتغير. وقد افترضنا أن $|a| < 1$ للسبب نفسه الذي افترضنا من أجله $|\gamma| < 1$.*

لاحظ أنه، في مرحلتنا هذه، طالما أن Y_t يعتمد على u_t في (6.60) فإن Y_t و u_t مرتبطان. وبالمثل، طالما أن Y_{t-1} تعتمد، بدورها، على u_{t-1} (من خلال الصيغة المبطأة لـ (6.60) فإنه يمكن إثبات أن هذين المتغيرين مرتبطان أيضا. وأخيرا، ينتج عن ذلك - بديهيا، في الأقل- أنه إذا كانت $\gamma \neq 0$ فإن Y_{t-1} تكون مرتبطة بـ u_t، لأنه من خلال (6.61) نجد أن u_t تعتمد على u_{t-1}، و Y_{t-1} مرتبطة بـ u_{t-1}.

إذا قمنا بتقدير (6.60) باستخدام منهج المتغير المساعد، فإن واحدة من المعادلات الطبيعية وهي $\Sigma_{t=1}^{n}(u_{t-1}\hat{u}_t) = 0$ سوف تناظر الافتراض $\text{cov}(Y_{t-1}, u_t) = 0$ ولكننا نلاحظ أنه إذا كانت $\gamma \neq 0$ فإن $\text{cov}(Y_{t-1}, u_t) \neq 0$، لذلك إذا ما استخدمنا $\Sigma(Y_{t-1}\hat{u}_t) = 0$ بوصفها إحدى هذه المعادلات الطبيعية، فإن مقدراتنا سوف تكون متحيزة. ويمكن إثبات أنها ستكون، أيضا، غير متسقة. والمشكلة الرئيسية في كل هذا هي أن ذلك التحيز وعدم الاتساق يجعل إحصائية d (للديربان واتسون) قريبة من 2 حتى ولو كانت $\gamma \neq 0$. تذكر أن احصائية d تقترب من 2 إذا كانت $\gamma \neq 0$. تذكر، أيضا، أن قيمة d القريبة من 2 تؤدي إلى قبول الافتراض

---

* نحتاج هذا الافتراض لجعل النموذج مستقراً. انظر ملحق هذا الفصل لفهم دور هذا الافتراض.

بعدم وجود ارتباط ذاتي $(\gamma = 0)$. ودلالة كل هذا هي أن استخدام اختبار ديربن واتسون في حالات تحتوي على متغيرات تابعة متباطئة، سيؤدي بالباحث إلى قبول الافتراض بعدم وجود ارتباط ذاتي بغض النظر عن وجود ذلك الارتباط الذاتي أو عدم وجوده. ومن الواضح أنه ينبغي ألا يستخدم اختبار ديربن واتسون في هذه الحالات.

ولحسن الحظ، يوجد اختبار آخر للارتباط الذاتي في حالة وجود متغير تابع متباطئ. ويتضمن هذا الاختبار استخدام إحصائية h لديربن.* افترض أنه تم تقدير (6.60) بوساطة طريقة المتغير المساعد التقليدية، وأن γ قد قدرت بوساطة (6.45). دع $\hat{\gamma}$ و $\hat{a}$ مقدرات لـ γ و a تم الحصول عليها، دع $\text{var}(\hat{a})$ أيضا، مقدرا لتباين $\hat{a}$ وأنه قد حصل عليه، حينئذ فإن إحصائية h تكون:

$$h = \hat{\gamma}\left(\frac{n}{1 - n\,\hat{\text{var}}(\hat{a})}\right)^{1/2} \tag{6.62}$$

حيث n حجم العينة، يمكن إثبات أنه إذا لم يكن هناك ارتباط ذاتي (أي $\gamma=0$) فإن h سوف تكون موزعة توزيعا طبيعيا معياريا $N(0,1)$ إذا كان حجم العينة لانهائيا. أما إذا كان هناك ارتباط ذاتي فإن h تصبح كبيرة وبالتحديد إذا كانت $\gamma > 0$ فإن h ستصبح كبيرة في الاتجاه الموجب. وعلى العكس إذا كانت $\gamma < 0$ تصبح h كبيرة ولكن في الاتجاه السالب.

توحي الملاحظات السابقة بالاختبار التالي لوجود الارتباط الذاتي في حالة وجود المتغيرات التابعة المبطأة بوصفها متغيرات مستقلة. هذا الاختبار هو اختبار للعينات الكبيرة، أو بمعنى آخر تكون النتائج صحيحة أو دقيقة، فقط، إذا كانت العينة ذات حجم لانهائي.

---

* يرجع إلى:

J. Durbin, "Testing for Serial Correlation in Least Squares Regression When Some of the Regressors are Lagged Dependent Variables," Econometrica 38 (1970), pp. 410-421.

اعتبر النموذج في (6.60)، وفرضية العدم لـ $H_0 : \gamma = 0$. دع الفرضية البديلة $H_1 : \gamma \neq 0$. حينئذ، يمكننا اختبار H_0 مقابل H_1 عند مستوى معنوية 0.05 بالطريقة التالية:

١- قدر (6.60) بالطريقة العادية مستخدما طريقة المتغير المساعد من الفصل الرابع ولاحظ قيمة $\hat{var}(\hat{\gamma})$ التي حصل عليها.

٢- من البواقي، تحسب $\hat{\gamma}$ كما في (6.45). وبالمقابل إذا كان البرنامج يزودنا باحصائية d لديربن واتسون فإنه يمكننا استخدام التقريب $\hat{\gamma} = 1 - d/2$. ويبني هذا التقريب على (6.54).

٣- نحسب قيمة إحصائية h لديربن.

٤- نرفض H_0 إذا كانت $|h| > 1.96$ ونقبلها إذا كانت h غير ذلك.

هناك تعديلان على هذا الاختبار (ينبغي أن يكونا واضحين). الأول إذا كان $H_1 : \gamma > 0$ واحتفظنا بمستوى 0.05 باعتباره مستوى معنوية فإننا سنرفض H_0 إذا كانت $h > 1.645$، والثاني إذا كان $H_1 : \gamma < 0$ فسنرفض H_0 إذا كانت $h < -1.645$ أما الاختبارات عند مستويات معنوية أخرى (مثلا عن 0.01) فهي تطبيقات مباشرة ونتركها للقارئ على سبيل التدريب.

هناك حدود على الاختبار المبني على احصائية h، فعلى سبيل المثال، وفي ضوء (6.62) ينبغي أن يكون واضحا أن الاختبار يفشل إذا:

$$n\,\hat{var}(\hat{a}) \geq 1 \qquad (6.63)$$

وذلك نظرا لأن h ستتضمن الجذر التربيعي لرقم سالب. في مثل هذه الحالات لايمكن تعريف h. لاتمثل هذه الحالة مشكلة من الناحية النظرية فحسب بل تحدث غالبا، في التطبيق، وفي مثل هذه الحالات، يمكن استخدام اختبار بديل. وهذا الاختبار هو اختبار للعينات الكبيرة، ولذا، تعد نتائجة تقريبية من الناحية التطبيقية.

مرة أخرى، دع فرضية العدم $H_0 : \gamma = 0$ والفرضية البديلة $H_1 : \gamma \neq 0$، بعدئذ، إذا أخذنا مستوى المعنوية عند 0.05، فإن خطوات الاختبار المقترح هي:

١- نقدر المعادلة الأساسية، مثلا (6.60)، بطريقة المتغير المساعد التقليدية.

٢- نحصل، بعد ذلك، على الأخطاء العشوائية المقدرة $\hat{u}_t$.

٣- نقدر باستخدام النتائج التي توصل إليها من الخطوة ٢ معادلة الانحدار التالية:

$$\hat{u}_t = a_0 + a_1\hat{u}_{t-1} + a_2 Y_{t-1} + c_1 X_{1t} + \cdots + c_k X_{kt} + \varepsilon_t \quad (6.64)$$

بالطريقة العادية. لاحظ أن (6.64) تحتوي على الخطأ العشوائي المقدر باعتباره متغيرا مستقلا وعلى كافة المتغيرات المستقلة في النموذج الأصلي (6.60).

٤- نحصل على نسبة t المناظرة لـ a_1 (أي $\hat{a}_1 / \hat{\sigma}_{\hat{a}_1}$). يمكن إثبات أنه $\hat{a}_1 / \hat{\sigma}_{\hat{a}_1}$ سيكون لها توزيع N(0,1) إذا كانت $\gamma = 0$ و n لانهائية.

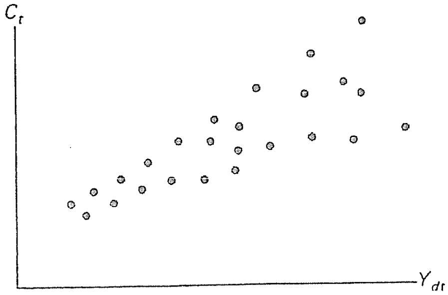
وبناء على ذلك، يتضمن الاختبار رفض H_0 إذا كان $\hat{a}_1 / \hat{\sigma}_{\hat{a}_1} > 1.96$ وقبولها في الحالات الأخرى. فإذا كانت الفرضية البديلة هي $H_1 : \gamma > 0$ فسنرفض H_0 إذا كان $\hat{a}_1 / \hat{\sigma}_{\hat{a}_1} > 1.645$. وتركت الحال التي يكون فيها الفرض البديل $H_1 : \gamma > 0$ على سبيل التدريب للقارئ.

(٦-٣) اختلاف التباين

نتناول في هذا المبحث مشكلة أخرى تنشأ نتيجة انتهاك أحد الافتراضات المرتبطة بالأخطاء العشوائية. تذكر أننا افترضنا في نموذجنا الأساسي للانحدار أن:

$$\text{var}(u_t) = E\left(u_t^2\right) = \sigma_u^2$$

أي أننا افترضنا أن جميع الأخطاء العشوائية لها التباين نفسه، وتعرف هذه الحالة (اصطلاحيا) بثبات التباين homoscedasticity. ولكن، قد يختلف تباين الأخطاء العشوائية، وفي هذه الحالة يطلق على هذه الحالة إختلاف التباين. وعلى سبيل المثال، قد نجد من دراسة لمستويات الإنفاق الاستهلاكي لأسر ذات دخول متاحة مختلفة، أن التباين في الاستهلاك يزداد مع ازدياد مستوى الدخل. فالأسر ذات الدخل المرتفع، مثلاً، قد تتميز بمرونة أكبر في الاستهلاك. ويظهر هذا الشرط بوضوح في الشكل رقم (٦-٦) حيث نجد أن مدى التغير للمجموعة الافتراضية من النقاط يتزايد عند مستويات الدخل الأعلى، وفي حالة مثل هذه، قد نفترض أن الخطأ العشوائي في دالة الاستهلاك يتسم باختلاف التباين.



شكل رقم (٦-٦)

نموذج أساسي

افترض أن لدينا دالة للاستهلاك تأخذ الشكل

$$C_t = b_0 + b_1 Y_{dt} + b_2 A_t + u_t, \qquad (6.65)$$

حيث:

C_t = الانفاق الاستهلاكي للأسرة t،

Y_t = الدخل المتاح للأسرة t،

A_t = الاصول السائلة التي تمتلكها الأسرة t، وأخيراً

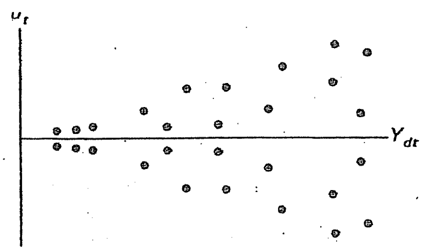
u_t = الخطأ العشوائي.

نفترض الآن أنه لأي مجموعة من قيم المتغيرات المستقلة، فإن u_t موزع توزيعا طبيعياً، وغير مرتبط ذاتيا لكن تباينه يرتبط بشكل متناسب مع دخل الأسرة t، أي $\mathrm{var}(u_t) = Y_{dt}\,\sigma^2_{u_t}$. وهكذا، كلما كان الدخل كبيرا نتوقع أن نشاهد تباينا أكبر في الاستهلاك.

ولقد افترضنا في نماذجنا السابقة أن الخطأ العشوائي مستقل عن جميع المتغيرات المستقلة، ولكن، ليس بوسعنا الآن أن نحتفظ بهذا الافتراض، لأننا حددنا أن حجم التباين يعتمد على قيمة أحد المتغيرات المستقلة Y_{dt}. لذلك، لم يعد الخطأ العشوائي مستقلا عن ذلك المتغير المستقل. وبدون وضع فروض إضافية، ليس بوسعنا أن نستخدم منهجنا العادي في التقدير، ذلك لأن التغاير بين الخطأ العشوائي والمتغير المستقل Y_{dt} لايساوي الصفر. وكما سنبين فيما

بعد، فإن الافتراض الإضافي الذي يمكننا من معالجة مشكلة اختلاف التباين ومن افتراض أن التغاير بين الخطأ العشوائي وبين المتغير المستقل Y_{dt} يساوي الصفر، وكيفما كانت قيم المتغيرات المستقلة Y_{dt} و A لأي مشاهدة هو أن يكون المتوسط الحسابي للخطأ العشوائي مساويا الصفر. واصطلاحا فإنه بالنسبة لأي قيم معطاة لكل من A_s, Y_{ds} تكون $E(u_t)=0$ لجميع t و s. وبالنسبة للمعادلة (6.65)، فهذا يدل على أن القيمة المتوقعة لـ C_t (C_t^m) لاتزال هي: $C_t^m = b_0 + b_1 Y_{dt} + b_2 A_t$.

ويدل هذا الافتراض، أيضا، على أن الخطأ العشوائي غير مرتبط بأي من المتغيرات المستقلة، أي $\text{cov}(u_t, Y_{dt}) = \text{cov}(u_t, A_t) = 0$. فإذا كان u_t مرتبط مثلا بـY_{dt} فإنه يتوقع أن تزداد قيمته أو تنقص مع زيادة Y_{dt} . ولكن افتراضنا بأن القيمة المتوسطة لـu_t تساوي الصفر لأي قيمة من قيم Y_{dt} يتضمن أن ذلك ليس صحيحاً، حيث إنه، مع زيادة Y_{dt}، تظل القيمة المتوقعة لـu_t ثابتة، أي صفراً. ينتج عن ذلك أن u_t وY_{dt} غير مرتبطتين.

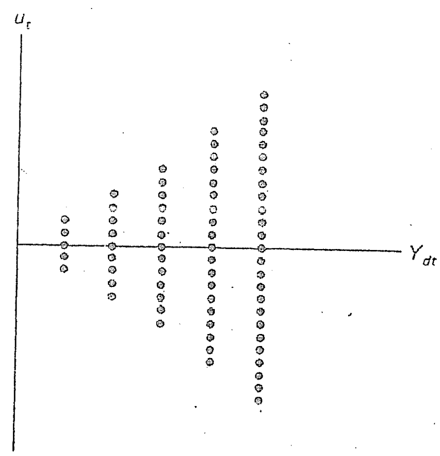


شكل رقم (٦-٧)

قد تبدو هذه النتيجة محيرة في ضوء افتراضنا بأن تباين u_t يزداد مع Y_{dt} . ولكن تزول هذه الحيرة بالنظر إلى الشكل (٦-٧) حيث يظهر من مجموعة النقاط الافتراضية أن تباين u_t يزداد مع تزايد Y_{dt}، ولكن من الواضح، أيضا، أن u_t لن يكون مرتبطا مع Y_{dt} لأن القيمة المتوسطة لـ u_t تساوي الصفر عند أي قيمة لـ Y_{dt} .

لاحظنا مما سبق أن u_t و Y_{dt} (وكذلك u_t و A_t) ليستا مرتبطتين، لأن القيمة المتوسطة لـ u_t تساوي الصفر عند أي قيمة معطاة لـ Y_{dt} (وكذلك A_t). فإذا كنا قد افترضنا أن قيمة u_t المتوسطة مساوية الصفر بدون التصريح بالشرط «لكل قيمة معطاة لـ Y_{dt}» فإنه لايمكننا الحصول على هذه النتيجة. افترض (مثلاً) أن X_1 متغير قيمته المتوسطة تساوي الصفر، $E(X_1) = 0$. دع $X_2 = 2\,X_1$ حينئذ تكون القيمة المتوسطة لـ X_2 مساوية للصفر $E(X_2) = 2\,E(X_1) = 0$، ولكن القيمة المتوسطة لـ X_2 لن تكون مساوية للصفر عند أي قيمة معطاة لـ X_1. على سبيل المثال إذا كانت $X_1 = 3$ فإن القيمة المتوسطة لـ X_2 تكون 6. في هذه الحال يكون X_1 و X_2 مرتبطان ببعضهما بصورة تامة.

وقبل أن نتجه إلى مشاكل التقدير علينا أن نوضح نقطة أخيرة قد تكون غامضة لكثير من القراء. نرى في الشكل رقم (٦-٨) وهو شكل منقح من الشكل رقم (٦-٧) أن جميع النقاط المناظرة لأي قيمة من Y_{dt} تظهر لها قيمة متوقعة مساوية الصفر. هذا يعكس أن القيمة المتوسطة لـ u_t تساوي الصفر لأي قيمة معطاة من Y_d.



شكل رقم (٦-٨)

لاحظ أن القيمة المتوسطة لجميع النقاط في الشكل تبدو مساوية الصفر، ويناظر هذا شرط أن القيمة المتوسطة (يطلق عليها، في بعض الأحيان، «المتوسط العام overall mean » لـ u هي الصفر. يتضح لنا الآن أن افتراض أن القيمة المتوسطة لـ u المناظرة لأي قيمة معطاة لـ Y_d المساوية الصفر تتضمن، بدورها، أن القيمة المتوسطة العامة مساوية للصفر أيضا. ولكن العكس ليس صحيحا بالضرورة، فقد تكون القيمة المتوسطة لـ u سالبة لبعضها قيم Y_d وموجبة لبعضها الآخر، ومع ذلك تظل القيمة المتوسطة العامة لـ u مساوية الصفر.

تأثيره على مقدراتنا

ماذا يحدث إذا استخدمنا المنهج العادي في التقدير لمعادلة تعاني مشكلة اختلاف التباين؟ بديهيا، يمكننا أن نتخيل أنه طالما أن القيمة المتوسطة للخطأ العشوائي لاتزال تساوي الصفر، وطالما أن u_t لاتزال غير مرتبطة بكل متغير من المتغيرات المستقلة [لذلك يكون $E(u_t\ Y_{dt}) = 0$ و $E(u_t\ A_t) = 0$ في المعادلة (6.65)] فإن مقدرات معلماتنا سوف تظل متسقة وغير متحيزة.* وأساسا فالنقطة المهمة هي أن شروطنا $E(u_t) = 0$، $E(u_t\ Y_{dt}) = 0$ و $E(u_t\ A_t) = 0$ لاتزال تفيد بأن $\Sigma\hat{u}_t = 0$ و $\Sigma(\hat{u}_t Y_{dt}) = 0$ وأخيرا $\Sigma(\hat{u}_t A_t) = 0$. فأساسا لايوجد خطأ في المعادلات الطبيعية. ولكن، كما هو الحال عند وجود الارتباط الذاتي، تكون هناك صيغ مختلفة لتباينات مقدرات معلماتنا. ومرة أخرى، إذا ما استخدمنا صيغنا العادية لتقدير هذه التباينات فستكون اختبارات الفرضيات وتكوين فترات الثقة الناتجة مشكوك فيها.

ينبغي أن يكون هذا واضحا، طالما أن نموذجنا الأساسي للانحدار يفترض

---

* يمكن للقارئ المهتم أن يثبت أن مقدراتنا المعتادة لاتزال غير متحيزة عن طريق العمل من خلال المناقشة الموجودة في ملحق الفصل الرابع. وعند القيام بذلك، علينا أن نلاحظ أن الافتراض الوحيد الذي نحتاجه في عملية الاشتقاق هو افتراض أن القيمة المتوسطة للأخطاء العشوائية هي الصفر لأي قيمة من قيم المتغيرات المستقلة وينتج هذا الشرط -في النموذج المعتاد- من افتراض الاستقلال، وقد أخذنا بهذا الافتراض، صراحة، في نموذجنا الذي يعاني اختلاف التباين.

تباينا ثابتا، وأن منهجنا في التقدير ينتج مقدرا لهذا الثابت. لكن، مع وجود اختلاف التباين، فلن يظل تباين الخطأ العشوائي ثابتا، وإنما سيكون متغيرا. وهذا يعني أن مقدرنا المعتاد سيمثل في الحقيقة أحد أنواع المتوسطات للتباينات المختلفة للأخطاء العشوائية، مثل هذا المقدر تكون له أهمية محدودة، ولايسمح لنا - على سبيل المثال - ببناء فترات ثقة صحيحة (أو نسب t) لعلمات المعادلة. وكما هو الحال عند وجود الارتباط الذاتي يمكننا الحصول على مقدرات أفضل لعلماتنا (أي مقدرات لها تباينات أصغر) ويمكننا أن نوجد مقدرات لهذه التباينات عن طريق إدخال معلومات ترتبط بالخصائص الحقيقية للخطأ العشوائي في منهجنا للتقدير.

طريقة للتقدير

لفحص مشكلة اختلاف التباين بعمق، دعنا نعود إلى علاقة الاستهلاك في المعادلة (6.65)، حيث جعلنا $var(u) = Y_{dt}\,\sigma_u^2$. سنثبت الآن أنه إذا ماقسمنا حدود المعادلة (6.65) بوساطة $\sqrt{Y_{dt}}$ فإننا سنحصل على معادلة تتسم بثبات التباين للخطأ العشوائي وينتج عن تلك القسمة:

$$\frac{Ct}{\sqrt{Y_{dt}}} = b_0\left(\frac{1}{\sqrt{Y_{dt}}}\right) + b_1\sqrt{Y_{dt}} + b_2\left(\frac{A_t}{\sqrt{Y_{dt}}}\right) + u_t^*, \qquad (6.66)$$

حيث:

$$u_t^* = \frac{u_t}{\sqrt{Y_{dt}}}$$

ولما كانت القيمة المتوسطة لـ u_t هي الصفر لأي مستوى من مستويات Y_{dt} فسيكون لدينا:*

$$E(u_t^*)\,E\left(\frac{u_t}{\sqrt{Y_{dt}}}\right) = \left(\frac{1}{\sqrt{Y_{dt}}}\right)E(u_t) = 0. \qquad (6.67)$$

أما بالنسبة لتباين u_t^* فسيكون لدينا وفقا للافتراضات نفسها:

* ينبغي أن يكون معنى (6.67) واضحا. إذا اعطيت Y_{dt} على أنها 900، وأن $E(u_t^*) = E(u_t)/30 = 0$ طالما أن القيمة المتوسطة لـ u_t تساوي الصفر لأي قيمة من قيم Y_{dt}.

$$\operatorname{var}(u_t^*) = E(u_t^*)^2 = E\left(\frac{u_t^2}{Y_{dt}}\right) = \frac{1}{Y_{dt}} E(u_t^2)$$
$$= \frac{1}{Y_{dt}} (Y_{dt}) \sigma_u^2 = \sigma_u^2 . \tag{6.68}$$

لنجد أن نموذجنا المعـدل (6.66) هو نموذج يتسم فيه الخطأ العشـوائـي، u_t^* بقيمة متوقعة صفرية وتباين ثابت.

وبالاستمرار في تحليلنا يتبين لنا أنه ليس من الصعب اكتشـاف أن u_t^* غير مرتبطة بالمتغيرات المستقلة في (6.66)، فعلى سبيل المثال، بالنسبة لأي مجموعة من المتغيرات المستقلة في (6.66) أو على نحو آخر، لأي قيم معطاة للمتغيرات المستقلة يكون $E(u_t^*) = (1/\sqrt{Y_{dt}})\, E(u_t) = 0$. يترتب على النتائج السابقة أن u_t^* غير مرتبـط بالمتغيرات المستقلة في (6.66) ولذلك نجد:

$$E\left[u_t^*\left(\frac{1}{\sqrt{Y_{dt}}}\right)\right] = 0, \tag{1}$$

$$E(u_t^* \sqrt{Y_{dt}}) = 0, \tag{2}$$

$$E\left[u_t^*\left(\frac{A_t}{\sqrt{Y_{dt}}}\right)\right] = 0. \tag{3}$$

باختصار، إذا قسمنا بالنسبة لكل مشاهدة كل من C_t ، Y_{dt} و A_t على $\sqrt{Y_{dt}}$ فإن نموذج الانحدار المناظر لهذه المجموعة الجديدة من القيمة الملاحظة «المصححة» سيصح (6.66)، ويحقق هذا النموذج جميع افتراضاتنا الأساسية. وحينئذ يمكننا ببساطة استخدام طريقتنا المعتادة في التقدير للحصول، في هذه الحالة، علـى مقدرات غير متحيزة لكل من معلمات الانحدار وتباينات المقدرات،* وخاصة

* على العكس من النماذج السابقة، لاتحتوي المعادلة (6.66) (التي تناظر المعادلات الطبيعية في 6.69) على حد ثابت، ويترتب على ذلك تغير في صيغ التباينات لمقدراتنا. وبالتحديد، فإن حدود في هذه الصيغ سـوف تعرف كبواقي انحدار (ولاتحتوي هذه على حد ثابت) المتغير المستقل رقم i على المتغيرات المستقلة الأخرى. على سبيل المثال، فإن تباين $\hat{b}_2$ في (6.69) سيكون $\sigma_u^2/\Sigma\hat{v}_{2t}^2$ حيث إن $\hat{v}_{2t}$ هو المتبقي من الانحدار:

$$\frac{A_t}{\sqrt{Y_{dt}}} = \gamma_1\left(\frac{1}{\sqrt{Y_{dt}}}\right) + \gamma_2\left(\sqrt{Y_{dt}}\right) + v_{2t}$$

فإنه باستخدام الافتراضات من (1) إلى (3) تكون معادلاتنا الطبيعية هي*

$$\sum\left(\frac{C_t}{Y_{dt}}\right)=\hat{b}_0\sum\left(\frac{1}{Y_{dt}}\right)+\hat{b}_1 n+\hat{b}_2\sum\left(\frac{A_t}{Y_{dt}}\right),$$

$$\sum C_t = n\hat{b}_0 + b_1\sum Y_{dt} + \hat{b}_2\sum A_t, \tag{6.69}$$

$$\sum\left(\frac{C_t A_t}{Y_{dt}}\right)=\hat{b}_0\sum\left(\frac{A_t}{Y_{dt}}\right)+\hat{b}_1\sum A_t+\hat{b}_2\sum\left(\frac{A_t^2}{Y_{dt}}\right).$$

ويزودنا المثال الذي شرحناه الآن، أيضا، بنظرة أعمق لمشكلة اختلاف التباين، فمنهجنا العادي للتقدير يعطي وزنا متساويا لكل مشاهدة من المشاهدات عند ايجاد مقدرات معلماتنا. بينما توضح مناقشتنا هذه أنه عند وجود مشكلة اختلاف التباين، فإنه ينبغي أن نعطي أوزانا مختلفة للمشاهدات، وبالتحديد، وعلى سبيل المثال ينبغي أن نعطي وزنا لكل مشاهدة يعادل($1/\sqrt{Y_{dt}}$) ويعني ذلك أنه ينبغي أن نعطي وزنا أقل للمشاهدات المناظرة للتباينات الكبيرة عن تلك المناظرة لتباينات أقل، وبديهيا يبدو هذا الأمر ذو معنى. فالمشاهدة التي تناظر تباينا أقل سوف تكون على الأرجح قريبة من خط الانحدار الحقيقي،** وفي منهجنا للتقدير ينبغي أن نعطي بعض الاهتمام لتلك النقاط التي نعتقد بأنها تقع أقرب إلى خط الانحدار الحقيقي عن تلك التي تكون في المتوسط بعيدة عنه، فالمشاهدات التي تناظر تباينات أقل هي ببساطة أكثر قيمة في تقدير موقع خط الانحدار عن تلك التي تكون انحرافاتها عن الخط أكبر.

---

* ينبغي أن نشير هنا إلى أن المعادلات الطبيعية آنفة الذكر قد اشتقت عن طريق وضع الافتراضات (1)-(3). وفي هذه الحالة، فإن الافتراض $E(u_t)=0$ لم يستخدم. أي أنه على الرغم من وجود ثلاث معلمات فقط، b_0 ، b_1 وb_2 فإنه يكون لدينا افتراضات أربعة خاصة بالخطأ العشوائي. وتنشأ هذه المشكلة، التي تتمثل في وجود عدد كبير من الافتراضات عادة، عند حل مشكلة اختلاف التباين والحل هو (عمومًا) أن يتم القيام بنفس العمل الذي قمنا به في المتن، وهو أن نستخدم فقط تلك الافتراضات التي تناظر التغايرات للمتغيرات المستقلة في النموذج المعدل. ويتضمن ذلك أننا يجب أن نهمل الافتراض بأن القيمة المتوسطة للخطأ العشوائي في النموذج المعدل يساوي الصفر، وعلى الرغم من أن اثبات ذلك خارج عن نطاق هذا الكتاب، فإنه يمكننا أن نبين أنه إذا اتبعت الطريقة نفسها، فإن المقدرات الناتجة لـ b_0 ، b_1 وb_2 ستكون لها تباينات أصغر عما لو فرضنا القيمة المتوسطة المساوية للصفر واثنين (أي اثنين) من الافتراضات الثلاثة السابقة.

** نعني بخط الانحدار الحقيقي معادلة القيمة المتوسطة التي تربط المتغير التابع بالمتغيرات المستقلة.

اختلاف التباين: طرق إضافية للمعالجة

كيف تعرف أنه لديك مشكلة اختلاف التباين؟ وكيف تغير - عمومًا - من طريقتك في التقدير لمواجهتها؟ ليست هذه أسئلة يسهل حلها (ولسوء الحظ غالبا مايتم تجاهلها). إحدى الطرق المقنعة لاكتشاف المشكلة هو أن نختبر أولا العلاقة محل الدراسة لنرى ما إذا كان هناك أي سبب للاعتقاد بأن الخطأ العشوائي يتسم باختلاف التباين. وغالبا ماتكتشف مشكلة اختلاف التباين من تكوين النموذج ذاته. افترض، على سبيل المثال، أننا نهتم باختبار الفرضية بأن مستوى الأرباح π - بالدولارات يعتمد على حجم مؤسسة الأعمال (مشارا إليه بوساطة قيمة أصولها A). يمكننا أن نعبر عن هذه العلاقة على النحو:

$$\pi_t = b_0 + b_1 A_t + u_t, \qquad (6.70)$$

حيث تتوافر لدينا مشاهدات عن π و A لعدد n من مؤسسات الأعمال. في ظل هذا النموذج، يصعب علينا قبول أن الخطأ العشوائي له التباين نفسه. وبالتأكيد، فإن التباين في الارباح سيكون أكبر بين منشآت مثل جنرال موتورز General Motors وستاندرد أويل Standard Oil عن المتاجر المحلية لبيع المنتجات الغذائية أو الأجهزة المنزلية، ويرجع هذا، فقط، إلى الاختلاف الكبير في الحجم المطلق لأرباحها. والنقطة المهمة هنا هي أننا نتوقع وجود علاقة طردية قوية بين قيمة A_t وتباين u_t. وبهذه المناسبة قد يكون هناك بعض المتغيرات المستقلة الأخرى في المعادلة (6.70) ولكن مشكلة اختلاف التباين تركز عادة على العلاقة بين واحد من المتغيرات المستقلة، وتباين الخطأ العشوائي، وعلى أي حال، في حالة مثل هذه، فإن رجحان وجود اختلاف التباين يكون كبيرا جدا.

توجد طريقتان أساسيتان يمكننا أن نستخدم أيا منهما لحل مشكلة اختلاف التباين. أولا: يمكننا إعادة صياغة العلاقة بطريقة يمكن معها إزالة اختلاف التباين. أي أنه يمكننا النظر إلى مشكلة اختلاف التباين على أنها مشكلة صياغة غير جيدة للنموذج، وهنا يمكننا أن نحل المشكلة بطريقة كفء عن طريق بناء نموذج افضل. وفي ضوء مثالنا أعلاه، فقد يكون من الأفضل أن نختبر العلاقة بين معدل الأرباح، أي $\pi^* = \pi/A$ وحجم المنشأة بدلا من العلاقة بين π و A وحينئذ، يمكننا تقدير المعادلة:

$$\pi^*_{t.} = b_0 + b_1 A_t + u_t, \qquad (6.71)$$

وبدون وجود سبب معين لتوقع اختلاف معدلات الأرباح اختلافا كبيرا بين المنشآت الكبيرة أو الصغيرة، يمكننا، بثقة أكبر، أن نفرض تباينا ثابتا بين الأخطاء العشوائية. ثانياً: يمكننا أن نحاول تحديد نمط اختلاف التباين ودمج هذه المعلومة في منهجنا للتقدير، فمثلا يمكننا في حالة دالة الاستهلاك (6.65) التي اخترناها سابقا أن نفترض أننا نعرف هذا النمط: يتناسب تباين u_t مع مستوى الدخل. وقد عالجنا هذه المشكلة عن طريق قسمة الحدود في نموذج الانحدار بوساطة الجذر التربيعي للدخل. ولكن، في كثير من الحالات، قد لايكون لدينا سبب كاف لافتراض نمط معين. ففي نموذج الارباح - الأصول (6.70)، على سبيل المثال، هل يكون تباين الخطأ العشوائي متناسبا مع مستوى الأصول A_t أم مع A_t^2 أو مع أي دالة أخرى لـ A_t. قد لانعرف الاجابة مسبقا عن هذه الأسئلة.

إن تحديد نمط اختلاف التباين يعد مشكلة صعبة. وبعض الحلول المقترحة لهذه المشكلة خارج نطاق هذا الكتاب*. ولحسن الحظ، يوجد منهج مباشر يمكن أن يؤدي إلى نتائج مشجعة. ففي ظل توافر افتراضات معينة يمكننا هذا المنهج من اختبار وجود اختلاف التباين، وأيضا من تحديد نمطه وأكثر من ذلك، فإن هذا المنهج يعتمد على المادة العلمية التي عالجناها في الفصول السابقة، ويسهل هذا، بالطبع، من فهمه، كما يزودنا ذلك أيضا بمراجعة مفيدة للمادة العلمية السابقة.

افترض أننا نرغب في تقدير العلاقة:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t, \qquad (6.72)$$

وأننا نشك أن تباين u_t مرتبط بانتظام بقيمة X_{2t} مثلاً، أي أننا نعتقد أن:

$$\sigma^2_{u_t} = f(X_{2t}). \qquad (6.73)$$

فإذا عرفنا الشكل المحدد للدالة $f(X_{2t})$، فإنه يمكننا (كما سبق) أن نحل مشكلة اختلاف التباين ببساطة من خلال قسمة معادلتنا (6.72) بوساطة $\sqrt{f(X_{2t})}$، لأن تباين الخطأ العشوائي الناتج $u_t^* = u_t / f(X_{2t})$ سوف يكون ثابتا ويساوي الواحد

الصحيح.* وسوف تحقق المعادلة الجديدة افتراض ثبات التباين ومن ثم، يمكننا إكمال التحليل على النحو المعتاد.

والمشكلة التي نواجهها هي أن الدالة $f(X_{2t})$ ليست معلومة. وسيكون طريقنا للحل أن نقول أولا بتقريب وتقدير $f(X_{2t})$ وبعد ذلك كما افترضنا من قبل - نقوم بقسمة (6.72) بوساطة الجذر التربيعي لمقدرنا $f(X_{2t})$. حينئذ يمكننا أن نستخدم معادلة الانحدار المعدلة لاشتقاق مجموعة من المعادلات الطبيعية التي يمكن أن نحلها للحصول على مقدرات معلماتنا.

لاحظ أولا أن افتراضنا بوجود اختلاف التباين في (6.75)، يتضمن أنه لقيمة معينة لـX_{2t}، أن:

$$E(u_t^2)=f(X_{2t}). \tag{6.74}$$

افترض الآن:

$$\varepsilon_t=u_t^2-f(X_{2t}) \tag{6.75}$$

ولاحظ أنه بالنسبة لقيمة معينة لـX_{2t}:

$$\begin{aligned}E(\varepsilon_t)&=E(u_t^2)-f(X_{2t})\\&=f(X_{2t})-f(X_{2t})=0.\end{aligned}$$

أي أن القيمة المتوسطة لـε_t هي الصفر.

دعنا الآن نحل (6.75) للحصول على u_t^2:

$$u_t^2=f(X_{2t})+\varepsilon_t. \tag{6.76}$$

لاحظ أن تفسير (6.76) واضح ومباشر، حيث تم التعبير عن u_t^2 كمجموع قيمته المتوسطة $f(X_{2t})$ ومتغير آخر ε_t يعكس انحرافه عن قيمته المتوسطة. لاحظ، أيضا، أن (6.76) يشبه كثيرا نموذج الانحدار.

---

* لاحظ أنه بالنسبة لقيمة معطاة من فإن:

$$E(u_t^*)=E\left[\frac{u_t^2}{f(X_{2t})}\right]=\frac{1}{f(X_{2t})}E\left(u_t^2\right)=\frac{f(X_{2t})}{f(X_{2t})}=1.$$

دعنا نفترض الآن أنه توجد لدينا مشاهدات عن u_t، وأن معلوماتنا المسبقة حول العلاقة تفيد أنه إذا كانت u_t تتسم باختلاف التباين كما في (6.74) فإن الدالة $f(X_{2t})$ قد يمكن، إلى حد ما، تقريبها بوساطة متعدد الحدود من الدرجة k.* يمكننا تحويل (6.76) في ظل هذه الافتراضات إلى نموذج انحدار يأخذ الشكل النمطي:

$$u_t^2 = a_0 + a_1 X_{2t} + \cdots + a_k X_{2t}^k + \varepsilon_t. \qquad (6.77)$$

لاحظنا أنه بالنسبة لأي قيمة معطاة من X_{2t} تكون $E(\varepsilon_t)$، كما لاحظنا أيضا أن هذا الشرط (القيمة المتوسطة الصفرية) تتضمن أن ε_t غير مرتبطة بـ X_{2t}. ولما كانت القيمة المعطاة لـ X_{2t} تتضمن أيضا، قيمة معطاة لكل قوى X_{2t}، فإنه ينتج عن ذلك أن ε_t يكون غير مرتبط أيضا، بكل واحدة من هذه القوى X_{2t}.** وهذا يعني أنه يمكننا أن نعامل (6.77) كنموذج انحدار له خطأ عشوائي ε_t يحقق الشروط كافة التي يمكن بوساطتها اشتقاق المعادلات الطبيعية.

افترض أننا نقوم بتقدير (6.77) بطريقتنا النمطية عن طريق جعل:

$$\sum \hat{\varepsilon}_t = 0, \sum (\hat{\varepsilon}_t X_{2t}) = 0, \cdots, \sum (\hat{\varepsilon}_t X_{2t}^k) = 0.$$

حينئذ يمكننا أن نثبت (باستخدام بعض الافتراضات الإضافية) - أن المقدرات الناتجة $\hat{a}_0, \hat{a}_1, \cdots, \hat{a}_k$ متسقة. وهكذا يكون المقدر المتسق لـ $f(X_{2t})$:

$$f(\hat{X}_{2t}) = \hat{a}_0 + \hat{a}_1 X_{2t} + \cdots + \hat{a}_k X_{2t}^k. \qquad (6.78)$$

والآن ينبغي أن يكون باقي هذا المنهج واضحا. حيث سنقسم نموذجنا الأولي (6.72) بوساطة $\hat{f}_t = \left[\hat{f}(X_{2t})\right]^{\frac{1}{2}}$، وبعدئذ، نحصل على مقدراتنا لـ b_0، b_1، b_2 من المعادلات الطبيعية:

---

* عادة تعد $k \geq 3$. وتعد نتائج هذا المبحث صحيحة، فقط، على المستوى النظري من التحليل إذا كان تقريب متعدد الحدود تاما. ولما كان يندر وقوع ذلك عمليا، فينبغي أن تعد جميع هذه النتائج تقريبية.

** على سبيل المثال، إذا كانت القيمة المتوسطة لـ ε_t هي الصفر على افتراض أن $X_{2t} = 3$ فإنها تكون مساويا الصفر أيضا عند $X_{2t}^2 = 9$.

$$\sum\left(\frac{Y_t}{\hat{f}_t}\right)=\hat{b}_0\sum\left(\frac{1}{\hat{f}_t}\right)+\hat{b}_1\sum\left(\frac{X_{1t}}{\hat{f}_t}\right)+\hat{b}_2\sum\left(\frac{X_{2t}}{\hat{f}_t}\right),$$

$$\sum\left(\frac{Y_tX_{1t}}{\hat{f}_t}\right)=\hat{b}_0\sum\left(\frac{X_{1t}}{\hat{f}_t}\right)+\hat{b}_1\sum\left(\frac{X_{1t}^2}{\hat{f}_t}\right)+\hat{b}_2\sum\left(\frac{X_{1t}X_{2t}}{\hat{f}_t}\right),$$

$$\sum\left(\frac{Y_tX_{2t}}{\hat{f}_t}\right)=\hat{b}_0\sum\left(\frac{X_{2t}}{\hat{f}_t}\right)+\hat{b}_1\sum\left(\frac{X_{1t}X_{2t}}{\hat{f}_t}\right)+\hat{b}_2\sum\left(\frac{X_{2t}^2}{\hat{f}_t}\right).$$

ولأن $\hat{f}_t$ مقدر متسق لـ f_t، فإنه يمكن إثبات (في ظل تحقق افتراضاتنا) أن المقدرات الناتجة تكون متسقة وكفئا. يضاف إلى ذلك أنه إذا كان حجم العينة لانهائيا، فإن صيغنا العادية للتباين تكون صحيحة. دع $\hat{\sigma}^2_{\hat{b}_i}$ حيث $i=0,1,2$ تشير إلى مقدر التباين $\hat{b}_i$ والذي نحصل عليه بوساطة صيغتنا العادية للتباين. حينئذ فإن اختبارات الفرضيات وتكوين فترات الثقة يمكن أن يتم انشاؤها عن طريق فرض أن $\left(\hat{b}_i-b_i\right)/\hat{\sigma}^2_{\hat{b}_i}$ متغير طبيعي معياري. وتكون هذه النتيجة (مرة أخرى) صحيحة، فقط، في حالة العينة اللانهائية. ولذلك ينبغي علينا عند التطبيق أن ننظر إلى النتائج على أنها تقريبية وبطريقة مشابهة لحالة نموذج الارتباط الذاتي يكون سبب هذا التعقيد هو أن مقدر $\hat{b}_i$ غير خطي في الخطأ العشوائي بسبب اعتماده على $\hat{f}_t$.

والصعوبة الواضحة في المنهج السابق هي أنه لن تتوافر لدينا مشاهدات عن u_t^2 ولذا، وقبل استخدام هذا المنهج، ينبغي علينا أولا أن نقدر قيم u_t^2. ويمكننا القيام بذلك بسهولة لأنه يمكننا أن نحصل على مقدرات متسقة للمعلمات، ومن ثم، للأخطاء العشوائية لنموذجنا الأصلي الذي يتسم باختلاف التباين (6.72) بوساطة طرقنا المعتادة في التقدير. حيث نقدر ببساطة معلمات النموذج (6.72) بالطرق المعتادة ثم نحصل بعد ذلك على مقدرنا للأخطاء العشوائية $\hat{u}_t=Y-\hat{b}_0-\hat{b}_1X_{1t}-\hat{b}_2X_{2t}$ ويمكن إثبات أنه إذا نفذ المنهج السابق بعد إحلال $\hat{u}_t^2$ محل u_t^2 فإن النتائج السابقة تظل صحيحة كافة.

اختبار لاختلاف التباين

تتوافر لدينا طريقة لتصحيح طريقتنا في التقدير في ظل وجود مشكلة اختلاف التباين للأخطاء العشوائية. ولإكمال هذا المبحث، دعنا نعود إلى قضية تحديد ما إذا كان نموذج الانحدار يعاني مشكلة اختلاف التباين أم لا. لقد ناقشنا من قبل (في نموذجنا المفترض لانحدار أرباح المنشآت على أصولها) كيف يمكننا أن نختبر معادلة الانحدار لمعرفة ما إذا كان تكوين النموذج نفسه يوصي برجحان وجود اختلاف التباين، ولكن، من المرغوب فيه، وجود منهج نظري لاكتشاف هل يوجد اختلاف التباين للأخطاء العشوائية. سوف نقدم الآن مثل هذا الاختبار. والفرضية التي سنقوم باختبارها هي أن الأخطاء العشوائية في نموذجنا الأولى (6.72) تتسم باختلاف التباين من النوع الموجود في (6.73). ينبغي علينا أن نعرف أن (6.73) تعني، رياضيا، أنه «إذا كانت الأخطاء العشوائية u_t تتسم باختلاف التباين، فإن نوع اختلاف تباينها سيكون كالمعطي في (6.73).

سوف نستخدم الآن تقريبنا لمتعدد الحدود(X_{2t}) لتكوين هذا الاختبار. وبالتحديد، يمكننا أن نجري اختبار النموذج في (6.73) عن طريق اختبار الافتراض المشترك في (6.77) أي $H_0 : a_1 = a_2 = \cdots = a_k = 0$ ، فإذا تم قبول H_0 فسوف نستنتج أن تباين u_t لايعتمد على(X_{2t})، وحينئذ سنعتبر u_t يتسم بثبات التباين. أما إذا رفضنا H_0 فسوف نستنتج أن u_t تتسم باختلاف التباين ولذا سوف نواصل من خلال منهج التقدير السابق.*

ويمكننا بناء اختبار عينة كبيرة للافتراض H_0 من خلال إحداث تغير في المنهج المقترح من الملحق (ب) (B) للفصل الخامس. وبالتحديد (مرة أخرى) دع $\hat{u}_t$ الخطأ العشوائي المقدر رقم t للنموذج (6.72) الذي يمكن الحصول عليه عن طريق تطبيق منهجنا المعتاد لذلك النموذج. دع ESS_u مجموع مربعات الخطأ الذي حصل عليه

* ينبغي علينا – نظرياً – قبل المضي في تصحيح مشكلة اختلاف التباين أن نحصل على عينة جديدة من المشاهدات. ولكن، في حالات عديدة، لايمكننا القيام بذلك، ولذا فسوف نستمر في العمل بعينتنا الأولية.

من خلال تكوين انحدار $\hat{u}_t^2$ على عدد (k+1) من المتغيرات في (6.77). أي الحد الثابت وX_{2t}^k ,..., X_{2t}، اجعل $\hat{\sigma}_t^2$ أيضا، هو التقدير المناظر لتباين الخطأ العشوائي الذي حصل عليه بالطريقة العادية أي$ESS_u / (N-K-1)$، حيث إن N: حجم العينة. دع أخيراESS_R هو مجموع مربعات الخطأ الذي حصل عليه عن طريق تكوين انحدار لـ$\hat{u}_t^2$ على الحد الثابت فقط. في هذه الحالة يكون $ESS_R = \Sigma_{t=1}^N\left(\hat{u}_t^2 - A\right)^2$ حيث إن $A = \Sigma_{t=1}^N \hat{u}_t^2 / N$. يمكننا حينئذ أن نثبت أنه إذا كان $N = \infty$ فإن$\left(ESS_R - ESS_u\right)/ \hat{u}_t^2$ سيكون متغيرχ_k^2 بدرجات حرية قدرها k. وبطريقة مشابهة للمناقشة الموجودة في الملحق ب (B) من الفصل الخامس، فإنه يمكن إثبات أنه إذا كانت الفرضيةH_0 غير صحيحة فسيكون الخطأ العشوائي متسما باختلاف التباين من النوع الموجود في (6.77) كما أنESS_R سيميل للكبر بالنسبة لـ ESS_u. لذلك فإن قيما كبيرة للاحصائية χ^2 و$\left(ESS_R - ESS_u\right)/ \hat{\sigma}_t^2$ ، ستؤدي إلى رفض H_0. وإذا رمزنا للمتغير χ^2 chi-square بدرجات حرية عددها k بالرمز χ_k^2. حينئذ فإنه بافتراض مستوى معنوية للاختبار ٥٪، سوف تعرف القيم الكبيرة للإحصائية χ^2 بأنها جميع القيم التي تزيد عن $\chi_{k,\ 0.95}^2$ حيث إن احتمال$[(\chi_k^2 \le \chi_{k,0.95}^2) = 0.95]$ ويمكن الحصول على قيمة$\chi_{k,\ 0.95}^2$ من أي جدول للمتغير χ^2.

وبالطبع، وعند التطبيق، فإن عينتنا لن تكون ذات حجم لانهائي، لذلك، ينبغي اعتبار نتائج اختبارنا تقريبية. ونلاحظ (على سبيل معلومة جديرة بالاهتمام) أنه إذا كانت$N = \infty$ فإنه يمكن إثبات أن هذا الاختبار لـχ^2 يكون معادلا لاختبار F الذي يمكن بناءه عن طريق اتباع المنهج الموجود في الملحق ب B للفصل الخامس بالنسبة للمعادلة (6.77)وذلك بعد إحلال $\hat{u}_t^2$ محل u_t.

ونشير أخيرا إلى أنه، على الرغم من أننا قد أنشأنا اختبارا لاختلاف التباين بدلالة متغير مستقل واحد، فإنه من السهل تعميم هذا الاختبار ليشمل حالة كون عدد من المتغيرات المستقلة مصدرا لاختلاف التباين. افترض، على سبيل المثال، أن تباين الخطأ العشوائي في (6.72) يعتمد على كل منX_{1t} وX_{2t}، أي:

$$\sigma'_{u_t} = g(X_{1t}, X_{2t}). \qquad (6.79)$$

ومرة أخرى، على افتراض أن الدالة غير معلومة، فإن منهجنا أساسا سيكون مماثلا لما أوضحناه سابقا فيما عدا أن (6.77) سيتم إحلالها بمتعدد في كل من X_{1t} و X_{2t}.

وللتوضيح، افترض أن k=2، حينئذ، سيتم إحلال المعادلة التالية محل المعادلة (6.7):

$$u_t^2 = a_0 + a_1 X_{2t} + a_2 X_{2t}^2 + b_1 X_{1t} + b_2 X_{1t}^2 + c_1 X_{1t} X_{2t} + \varepsilon_t. \qquad (6.80)$$

سيرتبط اختبار ثبات التباين بفرضية العدم $H_0 : a_1 = a_2 = b_1 = b_2 = c_1 = 0$. وبالتحديد، دع ESS_u: مجموع مربعات الخطأ التي حصل عليها من انحدار $\hat{u}_t^2$ على الحد الثابت والمتغيرات X_{2t}^2, X_{2t} ، X_{1t}^2, X_{1t} و $X_{1t} X_{2t}$. ولما كان لهذا الانحدار ست معاملات فإننا سنقدر σ_ε^2 على النحو: $\hat{\sigma}_\varepsilon^2 = ESS_u / (n-6)$. دع ESS_R (مرة أخرى): مجموع مربعات الخطأ الناتج من انحدار $\hat{u}_t^2$ على الحد الثابت فقط. يمكن في هذه الحال، إثبات أنه إذا كان حجم العينة لانهائيا فإن $(ESS_R - ESS_u) / \hat{\sigma}_\varepsilon^2$ سيكون متغير χ^2 خمسة درجات حرية. لاحظ أن الذي يحدد ESS_u (في هذه الحالة) له خمس متغيرات مستقلة.* فإذا أخذنا مستوى المعنوية مساويا لـ 0.05 فسنرفض في هذه الحالة H_0 إذا تحققت $(ESS_R - ESS_u) / \sigma_\varepsilon^2 > \chi^2_{5,0.95}$. ومرة أخرى، فإن نتائج هذا الاختبار تعد نتائج تقريبية فقط طالما أن حجم العينة، عادة، محدود من الناحية التطبيقية.

دعنا نذكر بعض الملاحظات الختامية لاختبارات التباين. إذا كانت درجة متعدد الحدود المقرب لنمط اختلاف التباين صغيرة ($k \geq 2$ مثلاً) فإن حدود التفاعل interaction terms التي تتضمن كلا المتغيرين المستقلين معا سوف تضمن في الانحدار. على سبيل المثال، يظهر الحد $X_{1t} X_{2t}$ في المعادلة (6.80) ولكن إذا كانت k كبيرة ($k \geq 3$ مثلاً) فإننا لن نهتم بجميع الحدود المتضمنة لأكثر من متغير مستقل واحد،

* عموما، تكون درجات الحرية لمتغير χ^2 المناظرة لـ $(ESS_R - ESS_u)/\hat{\sigma}_\varepsilon^2$ مساوية لعدد المتغيرات المستقلة (غير متضمنة الحد الثابت) المحددة لـ ESS_u

وسبب ذلك هو أنه إذا كانت k كبيرة وكل حدود التفاعل الممكنة بين المتغيـرات المستقلة متضمنة في النموذج. فإن عدد المتغيرات المستقلة سيكون كبيرا جدا، وقد يؤدي ذلك إلى ظهور مشاكل الارتباط الخطي المتعدد.

أما إذا كان حجم العينة لانهائيا، ولكن، k، درجة متعدد الحدود، محدودة (مهما كانت كبيرة) حينئذ فإن كل التفاعلات الممكنة ينبغي، من حيث المبـدأ، الاهتمام بها بهدف جعل اختبار اختلاف التباين جيدا بقدر الإمكان. وتعني عبارة «جيدا بقدر الإمكان» أنه عند مستوى معين من الخطأ من النوع الأول ينخـفـض الخطأ من النوع الثاني إلى حده الأدنى. ولكن، من الناحية العملية، يكون حجم العينة، عادة، محدوداً، ولذلك لاتوجد -في هذه الحالة- نتائج محددة تشير إلى العدد الذي يجب أخذه من حدود التفاعل، وذلك لجعل الاختبار جيـدا بـقـدر الإمكان. والتوجيه الوحيد الذي يمكن أن نقدمه في هذا المجال يكون التالي: دع p العدد الكلي للمتغيرات المستقلة في النموذج الذي يحدد مجموع مربعات الخطأ والمشار إليه بالرمـز ESS_u، حينئذ، نقترح أن يكون عدد الحدود التي تختبر فـي متعدد الحدود حيث تتحقق غير المتساوية التاليـة $(n-P) \geq 25$ وهذا الاقتراح مبني على حدسنا فقط.

اختبار آخر لاختلاف التباين

اختبار جولدفيلد – وكوندات Goldfeld - Quandt

نناقش الآن اختبارا آخر لاختلاف التباين، وهذا الاختبار، في ظل تحقق شروط معينة يكون سهلا ومشجعا. افترض، بالتحديد، أننا نعتقد أن واحدا من المتغيرات المستقلة هو مصدر اختلاف التباين. افترض، أيضا، أن العلاقة بين هذا المتغير وتباين الخطأ العشوائي مضطـردة monotonic ونعني بذلك أن تباين الخطأ العشوائي إما أن يتزايد باتساق مع قيمة المتغير المستقل، أو يتناقص باتساق مـع قيمة المتغير المستقل. على سبيل المثال، للتوضيح اقترحنا في مناقشتنا السابقة في المعادلة (6.70) احتمال تزايد تباين حجم الأرباح مع زيادة قيمة أصول المنـشـأة.

وتمثل هذه إحدى حالات العلاقة المتزايدة باضطراد بين تباين الخطأ العشوائي وأحد المتغيرات المستقلة.

إذا كان الخطأ العشوائي مرتبطا باضطراد مع أحد المتغيرات المستقلة وكان هذا الخطأ العشوائي موزعا توزيعا طبيعيا، فإنه يمكننا اختبار وجود اختلاف التباين باستخدام اختبار (جولدفيلد - كوندات)* (أو ج - ك)، هذا الاختبار له خصائص مشجعة. فبالمقارنة مع اختبارنا للعينات الكبيرة في المبحث السابق، فإن اختبار (ج - ك) هو اختبار للعينات الصغيرة. ولذا، ليست هناك ضرورة لاعتباره اختبارا تقريبيا للعينات الأقل من اللانهائية.

افترض وجود نموذج الانحدار الخطي التالي:

$$Y_t = b_0 + b_1 X_{1t} + \cdots + b_h X_{ht} + \cdots + b_k X_{kt} + u_t, \qquad (6.81)$$

حيث X_h هو المتغير المستقل الذي نشك بأنه مصدر اختلاف التباين. افترض أن الخطأ العشوائي موزع توزيعا طبيعيا، وإذا وجد اختلاف التباين، فسببه ارتباط تباين الخطأ العشوائي باضطراد مع X_h. بافتراض تحقق هذه الافتراضات، يتبع منهج اختبار (ج - ك) الخطوات التالية:

١- رتب جميع المشاهدات وفقا لقيم X_h. فإذا كانت العلاقة المشكوك فيها بين X_h وتباين الخطأ العشوائي علاقة مضطردة وموجبة، حينئذ ينبغي أن تكون أول مشاهدة أعيد ترتيبها تناظر أصغر قيمة لـ X_h. وتناظر المشاهدة التالية ثاني أصغر مشاهدة لـX_h وهلم جرا. وعلى العكس إذا كانت العلاقة المشكوك فيها بين X_h وتباين الخطأ العشوائي علاقة مضطردة وسالبة، ينبغي أن تكون أول مشاهدة أعيد ترتيبها تناظر أكبر قيمة لـX_h، وتناظر المشاهدة الثانية ثاني أكبر مشاهدة وهلم جرا. إذا رتبت المشاهدات بهذه الطريقة، ووجد اختلاف التباين فإن تباين الأخطاء العشوائية j, i سوف يكون حيث $var(u_j) < var(u_i)$ إذا كانت $j < i$ على سبيل

---

* يرجع إلى:

S. M. Goldfeld and R.E. Quandt. "Some tests for Homo-scedasticity". *Journal of American Statistical Association*, vol. 60 (1965), pp. 539-547.

التوضيح لإعادة الترتيب، افترض النموذج التالي:

$$Y_t = a_0 + a_1 X_{1t} + \cdots + a_2 X_{2t} + u_t, \qquad t = 1,2,\cdots,6 \tag{6.82}$$

بافتراض أن $\text{var}(u_t)$ يزداد بزيادة قيمة X_{1t}، حينئذ إذا كانت العينة الأولية هي:

(6.83)

	Y	X_1	X_2
t=1	10	2	15
t=2	12	1	27
t=3	-1	10	0
t=4	5	9	5
t=5	3	27	1
t=6	0	5	10

فإن العينة المعاد ترتيبها هي:

(6.84)

Y	X_1	X_2
12	1	27
10	2	15
0	5	10
5	9	5
-1	10	0
3	27	1

أما إذا فرض أن تباين u_t يتناقص مع زيادة قيمة X_{1t}، فإن العينة المعاد ترتيبها ستكون على النحو التالي:

(6.85)

Y	X_1	X_2
3	27	1
-1	10	0
5	9	5
0	5	10
10	2	15
12	1	27

٢- احذف عدد d من المشاهدات من وسط العينة المعاد ترتيبها. وعلى الرغم من أن الرقم d هو رقم تحكمي، إلا أن أخذ $d = n/3$، بالتقريب، يعد قاعدة حسابية

معقولة حيث إن n هو حجم العينة الأولية، يضاف إلى ذلك أن d ينبغي أن تختار حيث تكون (n-d) عددا صحيحا زوجيا، على سبيل المثال، إذا كانت n=61 تكون d=21 لذلك يكون n-d=40 وهو رقم زوجي، فإذا تم ذلك فإن العينة الأولية ستجزأ إلى عينتين فرعيتين يحتوي كل منهما على (n-d)/2 من المشاهدات.

وللتوضيح، إذا كانت المعادلة (6.83) هي العينة الأولية والمعادلة (6.84) هي العينة المعاد ترتيبها، حينئذ، تكون d = 6/3=2، ولذلك، سيتم إسقاط المشاهدتين الوسيطتين في المعادلة (6.84). وتكون العينة الفرعية الأولى هي أول (n-d)/2 = (6-2)/2=2 مشاهدتين.

Y	X_1	X_2
12	1	27
10	2	15

(6.86)

وستكون العينة الفرعية الثانية:

Y	X_1	X_2
-1	10	0
3	27	1

(6.87)

٣- قدر معادلات انحدار منفصلة لكل من العينتين الفرعيتين.

٤- احسب مجموع مربعات الخطأ (ESS) لكل واحدة من معادلات الانحدار. دع ESS_1 هو ESS للعينة الفرعية الأولى وESS_2 ليكون ESS للعينة الفرعية الثانية. لاحظ أنه، إذا وجد اختلاف في التباين، تكون الأخطاء العشوائية للعينة الفرعية الأولى ذات تباين أقل من تباين العينة الفرعية الثانية.

٥- لتبسيط الرموز، دع p عدد معاملات الانحدار (على سبيل المثال $(P=K+1)$ في المعادلة (6.81) حينئذ يمكن(ESS_2/ESS_1) إثبات هي موزعة بالضبط باعتبارها F بدرجات حرية عددها (n-d-2p)/2 في كل من البسط والمقام. وباجراء الخطوات السابقة فإننا سنرفض فرضية العدمH_0، ثبات التباين للأخطاء العشوائية عند مستوى المعنوية المختار، إذا كانت (ESS_2/ESS_1) أكبر من القيمة الحرجة لتوزيع F كما توجد في جدول F.

وعلى سبيل التوضيح، دع $e = (n-d-2p)/2$، ودعنا نرمز إلى المتغير F الذي له درجات حرية عددها e في كل من البسط والمقام بالرمز $F_{e,e}$. دع $F_{e,e}^{0.95}$ هي قيمة F من الجدول حيث يكون احتمال $F_{e,e} \leq F_{e,e}^{0.95} = 0.95$. حينئذ إذا أخذنا مرة أخرى، مستوى المعنوية ليكون (0.95) فإننا سنرفض H_0 إذا كانت $ESS_2/ESS_1 > F_{e,e}^{0.95}$.

وهكذا، فإن اختبار (ج – ك) اختبار بسيط ومباشر. إضافة إلى ذلك، فإن منطق هذا الاختبار واضح. فبديهيا، إذا كان تباين الخطأ العشوائي مرتبطا، حقيقة، بالمتغير المستقل (المشكوك فيه)، فإنه ينبغي علينا توقع أن تكون مربعات الأخطاء العشوائية المقدرة أكبر في العينة الفرعية الثانية منه في الأولى، لذلك، فإن القيم الأكبر لـ (ESS_2/ESS_1) تؤدي إلى رفض فرضية العدم، أما إذا كانت الأخطاء العشوائية المقدرة لها تقريبا الحجم في العينتين الفرعيتين، حينئذ، تقرب النسبة (ESS_2/ESS_1) من الواحد الصحيح، وحينها سنقبل فرضية العدم (أي ثبات التباين للأخطاء العشوائية) لأن القيمة الحرجة لاختبار F تكون أكبر من الواحد الصحيحا عند مستوى المعنوية 0.05 أو 0.01 ولدرجات حرية متساوية لكل من البسط والمقام.

هناك نقطتان ختاميتان ينبغي ملاحظتهما. أولاهما: يكون اختبار (ج – ك) صحيحا حتى إذا لم تحذف مشاهدات من وسط العينة المعاد ترتيبها. ولكن، أوضحت التجارب أن إلغاء بعض المشاهدات يحسن الاختبار لأنه يؤدي إلى تخفيض حجم الخطأ من النوع الثاني. والنقطة المهمة هي أن المشاهدات المحذوفة تجعل القسم الأول والقسم الثاني من العينة المعاد ترتيبها أكثر اختلافاً، وهكذا يصبح من السهل اكتشاف الاختلافات في التباين للأخطاء العشوائية في العينتين الفرعيتين.

والنقطة الثانية هي أن مناقشتنا السابقة قد افترضت، ضمنيا، أن $p < (n-d)/2$ أما إذا كانت $p > (n-d)/2$، فإن الاختبار لايمكن تنفيذه لأنه لن تكون هناك مشاهدات كافية في كل قسم من العينة لتقدير معلمات نموذج الانحدار، ومن ثم، لتحديد ESS_1 و ESS_2 إذا كان $(n-d)/2=p$ يكون عدد المشاهدات في كل قسم من العينة مساويا لعدد المعلمات التي تقدر. في هذه الحال، يمكن إثبات أنه إذا لم يكن هناك تعدد خطي تام بين المتغيرات المستقلة فإن كلا من ESS_1 و ESS_2 سيساوي الصفر، وهكذا

لايمكن، مرة أخرى، إجراء الاختبار. وقد افترضنا أنه إذا كانت d تعادل $n/3$ وتبين حينئذ أن $p \leq (n-d)/2$، حينئذ لاينبغي إسقاط أي مشاهدة من منتصف العينة المعاد ترتيبها (أي بجعل d مساوية الصفر). فإذا تم ذلك وكانت n عددا زوجيا فسيكون هناك $n/2$ من المشاهدات في كل من العينتين المعاد ترتيب كل منهما، أما إذا كانت $p < n/2$ فإن الاختبار ينبغي أن يجري كما شرحنا من قبل. وفي هذه الحال، ينبغي رفض H_0 إذا كان مستوى المعنوية 0.05 و $ESS_2/ESS_1 > F^{0.95}_{(n-2p)/2,\ (n-2p)}$. فإذا كانت n رقما فرديا نقترح جعل $n_1 = (n+1)/2$ في العينة الفرعية الأولى و $n_2 = n-n_1$ للعينة الفرعية الثانية. وهنا ينبغي أن يجرى الاختبار كما في السابق باستثناء أننا سنرفض H_0 إذا كانت: $Ess_2 / Ess_1 > F^{0.95}_{((n_1-p)/2,\ (n_2-p)/2)}$، وأخيرا، إذا كانت $n/2 \leq p$ فإنه لاينبغي استخدام اختبار (ج - ك) للكشف عن اختلاف التباين.

بعض التعليقات حول اختباري اختلاف التباين

افترضنا، حتى الآن، اختبارين لاختلاف التباين. ولكن هناك اختبارات أخرى، وسبب ذلك هو أنه لايوجد حتى الآن نموذج واحد يصلح لكل الظروف المحتملة، ففي ظروف معينة، يكون من الملائم استخدام اختبار معين، بينما في ظروف أخرى، يكون من الملائم استخدام اختبار آخر. وينبغي على الباحث أن يكون قادرا على استخدام هذين الاختبارين اللذين افترضناهما للتطبيق في معظم الحالات.

ولمعرفة القضايا المتضمنة في هذه الاختبارات، تذكر أنه، مع تحقق الافتراضات السابقة، فإن اختبار (ج - ك) مفيد لكونه اختبار عينة صغيرة، فليس مبنيا على تقريب للحالة المتضمنة عينة لانهائية. لذلك، فإن نتائجه دقيقة، وليس تقريبية. على سبيل المثال، فإن استخدام اختبار (ج - ك) عند مستوى معنوية 0.05 له في الحقيقة، خطأ من النوع الأول مساو لـ 0.05. إضافة إلى ذلك فإن التجارب قد اظهرت - مع تحقيق الافتراضات اللازمة لاستخدام الاختبار - أن الخطأ من النوع الثاني لاختبار (ج - ك) صغير بشكل مقبول. لذلك، إذا تحققت الافتراضات سالفة الذكر، فإن اختبار (ج - ك) يعد اختبارا جيدا للاستخدام.

ولكن، قد لاتتحقق الافتراضات سالفة الذكر. فقد تكون الأخطاء العشوائية، مثلا، موزعة توزيعا غير طبيعي، حينئذ ينبغي أخذ نتائج اختبار (ج - ك) على أنها نتائج تقريبية. وبديهيا، تقترت النتائج من الحقيقة كلما اقترب توزيع الأخطاء العشوائية من التوزيع الطبيعي. وفي الحقيقة، يرغب الاقتصاديون عند التطبيق افتـراض أن الأخطاء العشوائية موزعة توزيعا طبيعيا، ولذا، لايكون هذا الافتراض محل اهتمام كبير من جانبهم، وبالطبع، فنحن نذكر، فقط، وجهة النظر هذه دون أن نتبناها.

والافتراض الأكثر خطورة للمنهج ذاته هو افتراض أنه إذا وجـد اخـتـلاف التباين فهو بسبب ارتباط تباين الخطأ العشوائي باضطراد بواحد من المتغيرات المستقلة. هذا الافتراض يمكن الباحث من إعادة ترتيب العينة بطريقة لن يتناقص معها تباين الخطأ العشوائي وهذا هو حجر الزاوية للاختبار. ومن الواضح، أنه إذا كان تباين الخطأ العشوائي مرتبطا بأكثر من متغير مستقل واحد، أو كان مرتبطا بمتيغر مستقل واحد ولكن بطريقة أخرى غير مضطردة، حينئذ، لايستطيع الباحث عمومًا أن يعيد ترتيب العينة كما شرحنا من قبل، ومن ثم، لن يتمكن من إجراء هذا الاختبار.

يمكن للباحث في مثل هذه الحالات أن يستخدم الاختبار الذي شرحناه في المبحث السابق. ذلك الاختبار ليس محددا بالحالات التي يكون فيها واحـد مـن المتغيرات المستقلة، فقط، هو السبب في ظهور مشكلة اختلاف التباين، كما أنـه لايتطلب أن تأخذ العلاقة بين المتغيرات المستقلة المتسببة في اختلاف التباين وتباين الأخطاء العشوائية الصورة المضطردة. وفي الحقيقة، يمكن أن يتناول هذا الاختبار انماطا معقدة جدا من اختلاف التباين، وذلك من خلال جعل درجة متعدد الحدود المقرب، مثلا، في المعادلة (6.73) كبيرة. نلاحظ، أخيرا أن ذلك الاختبار ليس مبنيا على افتراض كون الأخطاء العشوائية موزعة توزيعا طبيعيا.

ولكن هذا الاختبار له حدوده، أيضا، فهو اختبار للعينات الكبيرة، ولما كانت العينات التي يستخدمها الباحث محدودة، فإنه لن يكون متأكدا من خصائص الاختبار الذي يستخدمه - على سبيل المثال حجم الخطأ من النوع الأول ! وأيضا، يـكـون حجم الخطأ من النوع الثاني - في هذا الاختبار كبيرا إذا كان حجم العينة صغيرا.

اختلاف التباين: نتيجة التجميع

سوف ننهي معالجتنا لاختلاف التباين بحال واقعية ظهرت في مجال تقدير الطلب على القمح في الولايات المتحدة الأمريكيـة*. ويختلف مصدر اختـلاف التباين في هذه الحالة عن المصادر التي عالجناها من قبل. وبشكل خاص تبـدأ الدراسة بنوع معتاد من دوال الطلب:

$$Q_t = b_0 + b_1 P_t^w + b_2 P_t^g + b_3 Y_t + b_4 D_t + b_5 S_{1t} + b_6 S_{2t} + b_7 S_{4t} + u_t. \qquad (6.88)$$

في هذه المعادلة، يعتمد الطلب على القمح في الفتـرة t (أو Q_t) على الأسعـار الجارية للقمح(p_t^w) وأسعار الحبوب الأخرى(p_t^s)، ومتوسط دخل الفرد(Y_t)، وأخيرا، على أربعة متغيرات صورية. يأخذ المتغير الصوري الأول منهـا(D_t) (والذي يأخذ في الحسبان تكلفة شهادات تسويق معينة ينبغي أن يشتريها المتعاملون في الصناعات الغذائية المحلية خلال جزء من الفترة موضع الاهتمام، ويأخذ قيمة الواحد الصحيح خلال تلك الفترات التي تكون فيها الشهادات فعالة، وقيمة الصفر في الأوقـات الأخرى. أما باقي المتغيرات الصورية S_1, S_2, S_4، فهي متغيرات صورية موسمية للفصول: الأول والثاني والرابع من السنة الميلادية على التوالي.

ومصدر المشكلة في هذه الحالة هو طبيعة البيانات، حيـث تـقـدم وزارة الزراعة الأمريكية أرقاما كل فترة زمنية عن الكميات والأسعار للقمح والحبـوب الأخرى. وتتوافر البيانات عن هذه المتغيرات على أساس ربع سنوي بدءا من الربع الثاني من سنة ١٩٦٤م، ولكن البيانات المتاحة قبل هذا التاريخ والمرتبطة بالطلب على القمح تتوافر، فقط، على أساس نصف سنوي.

على سبيل المثال، فإن المشاهدة المتاحة عن الطلب قبل الربع الثانـي عـام ١٩٦٤م مرتبطة بالنصف الأول (أول فصلين ربع سنويين) لسنة ١٩٦٤م، والمشاهدة المتاحة قبل هذه مرتبطة بالنصف الثاني (آخر فصلين ربع سنويين) لسنـة ١٩٦٣م

* اشترك في هذه الدراسة واحد من المؤلفين، انظر:

David Bradford and Harry H. Kelejian, *A Quarterly Demand Model for Wheat* (Unpublished Manuscript, 1976).

وهلم جرا. هناك مشكلة واضحة إذا أردنا استخدام البيانات قبل الفصل الثاني لعام ١٩٦٤م وبعده لتقدير دالة طلب موحدة. كيف يمكننا استخدام البيانات نصف السنوية وربع السنوية لتقدير معادلة انحدار واحدة؟

باستخدام نموذج أكثر عمومية، افترض أن نموذج الانحدار الموضح للقيم ربع السنوية للمتغير التابع Y_t هو:

$$Y_t = a_0 + a_1 X_{1t} + \cdots + a_n X_{nt} + u_t, \qquad (6.89)$$

حيث يشير الرمز السفلي t إلى الفصول ربع السنوية من السنة الميلادية، وحيث إن الخطأ العشوائي u_t يحقق الافتراضات المعتادة كافة. افترض - بالتحديد - أن u_t مستقل عن جميع القيم الماضية والحالية والمستقبلية للمتغيرات المستقلة وغير مرتبط ذاتيا، وله قيمة متوقعة مساوية للصفر $E(u_t) = 0$، وأخيرا، تباين ثابت $E(u_t^2) = \sigma_u^2$.

افترض أن المشاهدات ربع السنوية عن المتغير التابع متاحة، فقط، للفترات $(t = T, T+1, T+2, ..., T+N)$. أما بالنسبة للفترات السابقة عن الفترة T، فإنه توجد لدينا، فقط، مشاهدات نصف سنوية ليست متداخلة $(Y_{T-2} + Y_{T-1})$، $(Y_{T-4} + Y_{T-3})$، ...، $(Y_{T-\varphi} + Y_{T-\varphi-1})$، حيث إن φ هو رقم صحيح فردي يشير إلى عدد المشاهدات نصف السنوية المتاحة - وسنناقش هذا أدناه. في المثال السابق، ترتبط $(Y_{T-2} + Y_{T-1})$ بالنصف الأول من عام ١٩٦٤م، بينما ترتبط $(Y_{T-4} + Y_{T-3})$ بالنصف الثاني من عام ١٩٦٣م وهلم جرا. وبالعودة إلى نموذجنا (6.89)، نفترض أن المشاهدات ربع السنوية تكون متاحة لكل متغير مستقل ولجميع الفترات موضع الاهتمام (في المثال السابق - مثلا - قبل الفصل الثالث لسنة ١٩٦٤م وبعده).

إذا كان (6.89) هو نموذج الانحدار الذي يفسر المتغير ربع السنوي Y_t، فإنه يترتب على ذلك أن نموذج الانحدار للمتغير نصف السنوي $(Y_{T-j} + Y_{T-j-1})$ يكون:

$$\begin{aligned}(Y_{T-j} + Y_{T-j-1}) &= 20a_0 + a_1(X_{1,T-j} + X_{1,T-j-1}) + \cdots + a_n(X_{n,T-j} + X_{n,T-j}) \\ &\quad + (u_{T-j} + u_{t-j-1}), \qquad j = 1, 3, 5, \cdots, \varphi,\end{aligned} \qquad (6.90)$$

حيث φ هو رقم فردي صحيح يحدد قيمته (كما سنوضح فيما بعد) عدد المشاهدات نصف السنوية المتاحة. وقد امكن الحصول على النمـوذج (6.90) من خلال جمع الجانب الأيمن من (6.89) للفصول ربع السنوية التي تناظر المتغير التابع، أي $T-j$ و $T-j-1$.

وفي ظل تحقق افتراضاتنا، يكون للخطأ العشوائي$(u_{T-j} + u_{T-j-1})$ في النموذج نصف السنوي (6.90) قيمة متوقعة صفرية، وغير مرتبط ذاتيـاً، طالما أن المكونات ربع السنوية ليست متداخلة، ومستقلة عن القيم الماضية والجارية والمستقبلية كافة، ولها تباين ثابت وعلى وجه التحديد:

$$E(u_{T-j}+u_{T-j-1})^2=2\sigma_u^2 . \tag{6.91}$$

لذلك، يحقق نموذجنـا (6.90) المرتبط بالمشاهدات نصف السنوية جـمـيـع الافتراضات المعتادة، ولكن النموذج ربع السـنـوي (6.90) يحقق، أيضا، جميع الافتراضات المعتادة، ويحتوي على المعلمات غير المعلومة نفسها. سوف نبين الآن أن هذين النموذجين يمكن أن يدمجا معا في نموذج واحد يحقق افتراضاتنا المعتادة كافة، ويمكن استخدام كل من البيانات ربع السنوية ونصف السنوية لتقدير معلماته.

لاحظ أن مشاهداتنا المتاحة عن المتغير التابع يمكن ترتيبها زمنيا من الأقـدم إلى الأحدث كمايلي:

$$\begin{aligned}&(Y_{T-\varphi}+Y_{T-\varphi+1}),(X_{T-\varphi+1}),\cdots,(X_{T-5}+X_{T-6}),(Y_{T-3}+Y_{t-4}),\\&\qquad(Y_{T-1}+Y_{T-2}),Y_T,Y_{T+1},\cdots,Y_{T-N}.\end{aligned} \tag{6.92}$$

ولاحظ أنه إذا كانت $\varphi=5$ فسوف يكون لدينا ثلاث مشاهدات نصف سنوية هي $\left(Y_{T-2}+Y_{T-1}\right)$، $\left(Y_{T-4}+Y_{T-3}\right)$، $\left(Y_{T-6}+Y_{T-5}\right)$. لاحـظ، أيـضـا، أن $3=(5+1)/2$. وعلى سبيل مثال آخر ينبغي أن يكون واضحا أنه إذا كـانـت $\varphi=3$ فسيكون لدينا مشاهدتان نصف سنويتين. لاحظ، مرة أخـرى، أن $2=(3+1)/2$. وعمومًا توضح هذه الأمثلة أن المشاهدات نصف السنوية لأي رقم صحيح فردي يكون $(\varphi+1)/2$. لذلك يكون العدد الإجمالي للمشاهدات الموجودة في المعادلة (6.92) هو $\{[(\varphi+1)/2]+[N+1]\}$.

دعنا الآن نرمز للمشاهدات $\{[N+1]+[(\varphi+1)/2]\}$ في المعادلة (6.92) بالرمز y_t حيث إن $t=1,2,\ldots,\{[N+1]+[(\varphi+1)/1]\}$ وأن. أي أن y_t تشير إلى أقدم مشاهدة في (6.92). أي $(Y_{T-\varphi}+Y_{T-\varphi-1})$، y_2 تشير إلى المشاهدة التالية وهلم جرا. وبطريقة مشابهة نجد أن عدد المشاهدات نصف السنوية لكل متغير مستقل في المعادلة (6.90)* يبلغ $(\varphi+1)/2$. إضافة إلى ذلك، فقد افترضنا أن المشاهدات عن كل متغير مستقل تكون متاحة لكل فترة من الفترات التي تتاح فيها بيانات ربع سنوية للمتغير التابع أي للفترات $T,T+1,\ldots,T+N$. دع X_{jt} حيث $\{[N+1]+[(\varphi+1)/2]\}$,...,1=t) يرمز إلى المشاهدات $\{[N+1]+[(\varphi+1)/2]\}$ نصف السنوية وربع السنوية للمتغير المستقل X_{jt} التي رتبت زمنيا ترتيبا يشبه الطرق الموجودة في المعادلة (6.92). حينئذ، تتضمن نتائجنا في المعادلة (6.89) والمعادلة (6.90) أن y_t مرتبطة بـ X_{1t}، X_{2t}، ...، X_{nt} على النحو التالي:

$$y_t=a_0X_{0t}+a_1X_{1t}+\cdots+a_nX_{nt}+v_t,$$
$$t=1,2,\cdots,\left(\frac{\varphi+1}{2}\right)+(N+1), \tag{6.93}$$

حيث $x_{0t}=2$ إذا كانت t تناظر مشاهدة نصف سنوية، أي أن $t\le[(\varphi+1)/2]$ و $x_{0t}=1$ فيما عدا ذلك. وحيث إن u_t هو الخطأ العشوائي الذي يحقق جميع الافتراضات النمطية فيما عدا أنه يتسم باختلاف التباين. وبالتحديد فإن $E(v_t^2)=2\sigma_u^2$ إذا كانت $t\le[(\varphi+1)/2]$، وأن $E(v_t^2)=2\sigma_u^2$ في غير ذلك من الحالات.

يمكن تحويل النموذج (6.93) إلى نموذج يحقق افتراضاتنا المعتادة كافة. بالتحديد، دع $d=\sqrt{2}$ إذا كانت $t\le[(\varphi+1)/2]$ و $d_t=1$ إذا كانت t غير ذلك. حينئذ، يمكننا، باتباع المنهج الذي سبق توضيحه في المناقشات السابقة. أن نلغي مشكلة اختلاف التباين من خلال قسمة (9.93) على d_t وينتج عن ذلك النموذج:

---

* بحكم أننا افترضنا توافر المشاهدات ربع السنوية للمتغيرات المستقلة، فإنه من السهل تكوين المشاهدات نصف السنوية لتلك المتغيرات.

$$\left(\frac{y_t}{d_t}\right)=a_0\left(\frac{x_{0t}}{d_t}\right)+a_1\left(\frac{x_{1t}}{d_t}\right)+\cdots+a_n\left(\frac{x_{nt}}{d_t}\right)+w_t,$$
$$t=1,2,\cdots,\left(\frac{\varphi+1}{2}\right)+(N+1), \quad (6.94)$$

حيث إن $v_t / d_t = w_t$ ومن الواضح أن $E(w_t^2) = \sigma_u^2$ لجميع t، $\{t=1,2,...,[(\varphi+1/2)]+[N+1]\}$

والآن يستوفي النموذج (6.94) افتراضاتنا المعتادة كافة، ويرتبط بجميع مشاهداتنا المتاحة نصف السنوية وربع السنوية عن المتغير التابع. ويمكن تقدير هذا النموذج بوساطة منهجنا المعتاد في التقدير. وبالتحديد، نجد أن المعادلات الطبيعية يمكن اشتقاقها بوساطة الشروط:

$$\sum \hat{w}_t\left(\frac{x_{0t}}{d_t}\right)=0,\cdots,\sum \hat{w}_t\left(\frac{x_{nt}}{d_t}\right)=0, \quad (6.95)$$

حيث يتم كل تجميع على مدى t، $\{t = 1,2,...,[(\varphi +1/2) + [N + 1]\}$، وحيث تكون $\hat{w}_t=(y_t/d_t)-\hat{a}_0(X_{0t}/d_t),\cdots,\hat{a}_t(X_{nt}/d_t)$، حيث $\hat{a}_n,\cdots,\hat{a}_0$ مقدرات a_n, ..., a_0 على الترتيب.

نلاحظ أنه عندما يحصل على المقدرات $\hat{a}_n,\cdots,\hat{a}_0$ فإنه يمكن تفسير y_t بوساطة النموذج على النحو التالي:

$$\left(\frac{\hat{y}_t}{d_t}\right)=\hat{a}_0\left(\frac{x_{0t}}{d_t}\right)+\cdots+\hat{a}_n\left(\frac{x_{nt}}{d_t}\right), \quad (6.96)$$

أو بالغاء d_t على النحو:

$$\hat{y}_t=\hat{a}_0x_{0t}+\cdots+\hat{a}_nx_{nt} \quad (6.97)$$

ويعني هذا أن احدث القيم ربع السنوية تفسر على النحو:

$$\hat{Y}_t=\hat{a}_0+\hat{a}_1X_{1t}+\cdots+\hat{a}_nX_{nt}, \qquad t=T,T+1,\cdots,T+N; \quad (6.98)$$

بينما تفسر القيم نصف السنوية على النحو:

$$(\hat{Y}_{T-j}+\hat{Y}_{T-j-1}) = 2\hat{a}_0+\hat{a}_1(X_{1,T-j}+X_{1,T-j-1}) \\ +\cdots+\hat{a}_n(X_{n,T-j}+X_{n,T-j-1}), \quad j=1,3,5,\cdots,\varphi.$$

(٦-٤) مشاكل في اختيار المتغيرات

افترضنا، حتى الآن، أن متغيرات نموذج الانحدار تعطي لنا بطريقة أو بأخرى، وأن مشكلتنا الوحيدة تنحصر في تقدير النموذج واختبار الفرضيات وعلاج الارتباط الذاتي وهلم جرا. ولكن، عند التطبيق، نجد أنه ينبغي علينا اختيار المتغيرات التي يشتمل عليها النموذج. ويعتمد الباحث في تحديده للمتغير التابع على إحدى النظريات، ثم يحاول، بعد ذلك، تحديده المتغيرات المستقلة التي تفسر النظرية تفسيرا أفضل. وعند القيام بذلك يمكن أن يقع الباحث في نوعين من الأخطاء. أولاً: قد يفشل الباحث في تضمين أحد المتغيرات المستقلة مهمة في نموذجه، بمعنى أنه قد يغفل أحد العوامل المهمة المحددة للمتغير التابع. ثانيا قد يفترض الباحث أن عاملا معينا يكون مهما في تحديد المتغير التابع في الوقت الذي لايكون فيه ذلك العامل مهما في الحقيقة. فإذا ماحدث ذلك فسيشتمل نموذجه على متغير غير ضروري. سنهتم، في هذا المبحث، بمناقشة نتائج الوقوع في كل واحد من هذه الأخطاء.

متغير محذوف

دعنا نهتم أولا بحالة حذف المتغيرات المستقلة من العلاقة المفترضة، افترض، على سبيل المثال، أن العلاقة الحقيقية (غير المعلومة لنا) هي:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t, \tag{6.99}$$

حيث u_t الخطأ العشوائي الذي يستوفي جميع افتراضاتنا المعتادة، ولكننا أغفلنا X_2 وأعتبرنا أن المعادلة تأخذ الشكل التالي:

$$Y_t = b_0 + b_1 X_{1t} + r_t, \tag{6.100}$$

حيث جعلنا r_t خطأنا العشوائي. سنبين الآن أنه إذا أردنا تقدير المعادلة (6.100) بطريقتنا العادية، فإن المقدرات الناتجة لكل من b_0 و b_1 سوف تكون، عمومًا، متحيزة وغير متسقة. والسبب هو أن الفشل في إدخال X_2 في النموذج يؤدي إلى خرق لافتراضاتنا الأساسية المرتبطة بالخطأ العشوائي، وهو في هذه الحالة، r_t.

وبالتحديد وبمقارنة المعادلة (6.99) مع المعادلة (6.100) سنرى أن الخطأ العشوائي - في النموذج الذي يجري تقديره - يعتمد جزئيا على X_{2t}.

$$r_t = b_2 X_{2t} + u_t \,, \tag{6.101}$$

وبأخذ القيم المتوقعة نجد:

$$E(r_t) = E(b_2 X_{2t}) + E(u_t) = b_2 \mu_2 + 0 = b_2 \mu_2, \tag{6.102}$$

حيث μ_2 هي القيمة المتوسطة لـ X_{2t}. ومن الواضح أن القيمة المتوقعة للخطأ r_t لن تكون، عموما، مساوية للصفر، فيما عدا الحالة التي تكون فيها $b_2 = 0$، والتي تعني أن Y_t لاتعتمد على X_{2t}.* وهكذا، نرى أنه إذا كانت Y دالة في X_2 في معادلة الانحدار، فإن الخطأ العشوائي في تلك المعادلة لن يكون، له عمومًا، قيمة متوسطة صفرية. يضاف إلى ذلك أنه ينبغي أن يكون واضحا ارتباط r_t بـ X_{1t} إذا كانت X_{2t} مرتبطة بـ X_{1t}، ونتيجة لذلك، فإن التغاير $\text{cov}(r_t, X_{1t})$ لن يكون، عمومًا صفرا. وينتج عن ذلك، في الأقل، بديهيا، في ظل مثل هذه الشروط أن مقدراتنا لـ b_0 و b_1 ستكون متحيزة وغير متسقة. فمثلا، إذا كانت $E(r_t) \neq 0$ ولكننا حصلنا على معادلتنا الطبيعية الأولى عن طريق وضع $\Sigma \hat{r} = 0$، فإن منهجنا في التقدير لن يتسم بالاتساق.

وقد يكون من المفيد أن نحدد طبيعة هذا التحيز عند مستوى تحليلي أعمق. افترض أننا نرغب في تقدير دالة استهلاك وأن العلاقة الحقيقية تأخذ الشكل:

$$C_t = b_0 + b_1 Y_{dt} + b_2 A_t + u_t, \tag{6.103}$$

* والحالة الأخرى المناظرة هي عندما يكون $E(r_t)=0$، وهي حالة عرضية وفريدة لأنها تعني أن $\mu_2 = 0$.

حيث:

C_t الانفاق الاستهلاكي خلال الفترة الزمنية t،

Y_{dt} الدخل المتاح خلال الفترة الزمنية t،

A_t حجم الأصول السائلة في الفترة t،

u_t الخطأ العشوائي

نتوقع أن يكون لـA_t تأثير علىC_t، ولذلك، فإن $b_2 > 0$. افترض، أيضا، أنه توجد علاقة ارتباط موجب بـينA_t وY_t، بمعنى أنه كلما تزايد الدخل المتاح تتزايد أيضا، قيمة الأصول السائلة (والعكس صحيح). ولكن افترض أن المعادلة التي نقدرها فعلا لاتحتوي علىA_t:

$$C_t = b_0 + b_1 Y_{dt} + r_t. \tag{6.104}$$

لاتكون القيمة المتوقعة للخطأ العشوائي في هذه الحالة (كما في التحليل السابق) للمعادلة التي اخترنا تقديرها مساوية للصفر، عمومًا:

$$E(r_t) = E(b_2 A_t) + E(u_t) = b_2 \mu_A \neq 0, \tag{6.105}$$

حيث إنμ_A هي القيمة المتوسطة لــ A_t. وبالمثل، يمكن أن نتوقع أن يرتبطr_t بـ Y_{dt}. وسوف نحصل، لذلك، على مقدرات متحيـزة لـــb_0 وb_1 (حيث يشيـر الأخير إلى م ح س). أكثر من ذلك، فإنه، بسبب التـأثـيـر المـوجـب لـــA_t علىC_t($b_2 > 0$) والارتباط الموجب بينA_t وY_{dt}، فسوف نحصل على تقدير أعلى مما يجب لـ م ح س، بمعنى أن $E(\hat{b}_1) > b_1$. وسبب ذلك، بديهيا، هو أنY_{dt} في المعادلة (6.104) تعمل باعتبارها متغيرا ينوب عن نفسه وعن A_t. أي أنه ينسب إليه التأثير الموجب لكل مـنY_{dt} وA_t على C_t. والنقطة المهمة هنا هي أنه، عندمـا يرتفعY_{dt} فإنA_t ترتفع كذلك ولكن غيـابA_t من المعادلة المقدرة ينسب التأثيـر الموجب لـA_t علىC_t إلىY_{dt}، ونتيجة لذلك، فإن التأثير المنسوب لـY_{dt} علىC_t ارتباطا سالبا ولكنها ماتزال مرتبطة بـY_{dt} ارتباطا موجبا، حينئذ، فإن مقدرنا لـ b_1 في غيابA_t من معادلة الانحدار سيكون متحيزا لأسفل، وسينعكس الأثر السالب لـA_t علىC_t، بعض الشئ، فـي $\hat{b}_1$. نلاحظ، أخيرا، أن اتجاه التحيز في هـذه

الحالات سيكون (عمومًا) معكوسا إذا كانت Y_{dt} و A_t مرتبطتين ببعضهما ارتباطا سالبا. والنقطة المهمة في كل هذا ليست هي تحديد اتجاه التحيز بقدر ماهي إدراك أنه إذا حذف أحد المتغيرات المهمة من معادلتنا فإن مقدراتنا سوف تكون متحيزة وغير متسقة. وحقيقة فإنه لايمكن، عادة، تحديد اتجاه التحيز في معظم نماذج الانحدار التي تشتمل على عدد كبير من المتغيرات المستقلة.

متغيرات أكثر من اللازم

دعنا الآن نهتم بحالة وجود عدد كبير من المتغيرات المستقلة في النموذج. افترض، على سبيل المثال، أن العلاقة الحقيقية تأخذ الشكل:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t , \qquad (6.106)$$

حيث نفترض أن الخطأ العشوائي مستقل عن قيم جميع المتغيرات المستقلة، وأنه يحقق الافتراضات المعتادة الأخرى كافة. تبين هذه المعادلة أنه ليس من الضروري أن ندخل المتغير X_{3t} في نموذجنا لأن Y_t لاتعتمد عليه أي أن:

$$b_3 = 0$$

لاحظ (مع ذلك) أن هذا لايمنعنا من كتابة نموذجنا على النحو:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + b_3 X_{3t} + u_t , \qquad (6.107)$$

حيث إن $b_3 = 0$. تذكر أننا في مناقشتنا للملحق ب (B) للفصل الخامس، قد أوضحنا أنه طالما أن الصفر هو رقم ثابت، فلايوجد، أساسا، خطأ في وجود معامل انحدار ذي قيمة مساوية للصفر.

افترض أننا لانعلم أن b_3 هي، في الحقيقة، مساوية للصفر. ونتيجة لذلك، نعد المعادلة (6.107) نموذجنا، ونكمل التحليل للحصول على مقدرات b_0، b_1، b_2 و b_3 بالطريقة العادية، حينئذ، يكون السؤال الذي ينبغي أن نحاول الاجابة عنه هو: هل ستعطينا الطريقة العادية في التقدير مقدرات غير متحيزة لمعاملات الانحدار ولتباين مقدراتنا، والإجابة هي، لحسن الحظ، نعم، طالما أنه لم ينتهك أي من الافتراضات الأساسية المرتبطة بالخطأ العشوائي في المعادلة (6.107). افترض، على

سبيل المثال، أن u_t مستقلة عن X_{3t}. حينئذ سيكون لدينا نموذج انحدار في المعادلة (6.107) تكون القيمة المتوسطة للخطأ العشوائي به مساوية للصفر، ويكون ذلك الخطأ العشوائي مستقلا عن المتغيرات المستقلة كافة كما يحقق جميع افتراضاتنا الأخرى. وينتج عن ذلك أن مقدرات b_0 , b_1 , b_2 و b_3 غير متحيزة.* وبديهيا إذا لم نعرف أن تأثير X_{3t} على Y_t هو الصفر فإن البيانات ستدلنا على ذلك، في شكل أن مقدرنا لـ b_3 سيكون غير متحيز $E(\hat{b}_3) = b_3 = 0$. وبالمقابل إذا أردنا اختبار فرضية العدم $b_3 = 0$ عند مستوى معنوية 5% فإنه ستكون لدينا فرصة ٩٥٪ لقبوله.

تعقيبات إضافية

قد يبدو من النتائج السابقة أننا لانخسر شيئا إذا ما أدخلنا مجموعة كبيرة من المتغيرات المستقلة المشكوك في أهميتها في المعادلة التي نرغب في تقديرها. ولكن هذا ليس صحيحا. لاحظ أولا أنه إذا كانت هذه التقديرات مستقلة عن الخطأ العشوائي حيث إنه لم ينتهك أي من افتراضاتنا فإنه لايزال بإمكاننا رفض فرضية العدم المتعلقة بما إذا كانت إحدى المعلمات المعنية تساوي الصفر أم لا. وبمعنى آخر، فقد نقع في الخطأ من النوع الأول.

ثانيا، وربما كان ذلك أكثر أهمية، يتزايد تباين مقدرات المعلمة - مع ثبات حجم العينة - عموما - مع تزايد عدد المتغيرات المستقلة. أي أنه إذا علمنا أن $b_3 = 0$ وهكذا حصلنا على مقدر مثلا لـ b_2 من المعادلة (6.106)، فإن هذا المقدر، عموما، سيكون له تباين أصغر من مقدر b_2 الذي نحصل عليه من المعادلة (6.107) ويعني هذا أننا باستخدام المعادلة (6.107) نطلب كثيرا من البيانات المتاحة، طالما أنه

---

* هذه هي الحالة العادية التي يهتم بها في محاضرات الاقتصادي القياسي، ويمكن، في الحقيقة، إثبات أنه، إذا كانت X_{3t} غير مرتبطة بالخطأ العشوائي فإن مقدراتنا - مع توافر افتراضات إضافية - تظل متسقة. وأخيرا، ينبغي أن يكون واضحا، أنه إذا كانت X_{3t} مرتبطة بالخطأ العشوائي، فإن مقدراتنا ستكون متحيزة وغير متسقة.

ينبغي أن نقدر أربع معلمات بدلا من ثلاث. وبالمقابل، يدمج نموذج (6.106)، على العكس من (6.107) المعلومة بأن $b_3 = 0$، ويعكس تباين المقدرات هذه المعلومة الإضافية. إذا ثمة تكلفة مرتبطة مع زيادة عدد المتغيرات المستقلة: تباينات أكبر تقود لمقدرات أقل دقة للمعاملات. ونتيجة لذلك، تصبح فترات الثقة أوسع، وقد نرفض أحد المتغيرات المستقلة في النموذج على أساس أنه غير مهم إحصائيا بينما له في الحقيقة تأثير منتظم على المتغير التابع.

ويوقعنا هذا في معضلة. إذا تركنا أحد المتغيرات نحصل على نتائج متحيزة، ولكن، من ناحية أخرى، إذا أدخلنا عددا كبيرا من المتغيرات في النموذج يزيد تباين مقدراتنا. ويشير هذا إلى الحاجة للاهتمام والاجتهاد عند اختيار مجموعة المتغيرات المستقلة. وحتى إذا كنا مهتمين، فقط، بالعلاقة بين متغيرين اثنين فقط (كما في المثال الموجود في الفصل الرابع حيث يهتم الباحث بتقدير تأثير معدل ضريبة المبيعات على مبيعات التجزئة في مراكز المدن)، فإنه من الضروري جعل النموذج يشتمل على متغيرات أخرى تحدد قيمة المتغير التابع. والفشل في ادخال المتغيرات المستقلة الأخرى سوف ينتج عنه، عمومًا، مقدر متحيز للمعامل الذي نهتم به. وبالعودة مرة أخرى إلى ذلك المثال، سوف يتجه التأثير المقدر لمعدل الضرائب الأعلى على حجم المبيعات بمركز المدينة للتحيز إذا لم ندخل المتغيرات الأخرى المحددة لمبيعات التجزئة في المدينة. ومن الناحية الأخرى، لايوجد سبب لادخال أي متغير تتوافر عنه البيانات باعتباره متغيرا مستقلاً في النموذج، ينبغي عليك أن تختار، فقط، تلك المتغيرات التي تعتقد على أسس مسبقة أنها تؤثر في المتغير التابع.

وأحد المناهج التي يستخدمها الاقتصاديون بانتظام هي أن «يجربوا» المتغير الذي يعتقدون بأهميته ثم يسقطونه من النموذج إذا تبين أنه ليس مختلفا عن الصفر معنوياً. افترض، على سبيل المثال، أننا نريد تقدير المعادلة:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t \,, \qquad (6.108)$$

حيث تتوافر لدينا، فقط، أسباب واهية للاعتقاد بأن X_2 يؤثر في Y. ووجدنا فعلا أنه لايمكننا رفض فرضية العدم بأن $b_2 = 0$. وأحد المناهج البديهية المغرية التي يمكن اللجوء إليها لتقليل تباين مقدرنا لـ b_1 هو اسقاط X_2 من المعادلة وتقدير المعادلة المنقحة:

$$Y_t = b_0 + b_1 X_{1t} + u_t \,, \qquad (6.109)$$

وبينما يكون لهذا الأسلوب الشائع الاستخدام قية عملية بإسقاطه للمتغيرات التي ليست لها أهمية واضحة، فإننا يجب أن نكون على حذر من أن هذه الطريقة التي تستخدم، عادة، تتضمن تعقيدات معينة. وعلى سبيل أحد الأمثلة إذا كانت المجموعة الأولية من البيانات «معادا استخدامها» لتقدير (6.109)، فإن عنصرا من الدائرية يدخل في النموذج، ويمكن اظهار أن المقدرات الناتجة تكون متحيزة. ويدخل هذا العنصر من الدائرية لأن النموذج (6.109) يعاد بناؤه على أساس الاختبار الذي أجرى على البيانات المستخدمة في تقدير النموذج. ومن الواضح أن الطريقة العلمية تتطلب من الباحثين أن يكونوا نماذجهم قبل تحليل البيانات التي ستستخدم لتقديرها.

هناك مشاكل أخرى ترتبط بهذه الطريقة المتابعة، وتظهر بعض هذه المشاكل حتى مع وجود مجموعة جديدة من البيانات تستخدم في تقدير (6.109). والمناقشة الكاملة لهذه المشاكل تقع خارج نطاق هذا الكتاب، ولكن ينبغي على القارئ أن يدرك أنه تنشأ مشاكل معينة كلما استخدمت نتيجة اختبار، في اختيار النموذج الذي يستخدم لتقدير المعلمة. وهذا مافعلناه بالضبط عندما قدرنا b_1 بدلالة (6.108) إذا كان X_2 احصائيا معنويا، وبدلالة (6.109) إذا كان X_2 غير معنوي. في مثل هذه الحالات، لاتحتفظ المقدرات بخصائصها المعتادة، فعلى سبيل المثال، قد تكون المقدرات متحيزة، وتبايناتها قد لايمكن الحصول عليها بوساطة الصيغ النمطية، وقد لاتصبح موزعة توزيعا طبيعيا. وعلى الرغم من أن هذه الطريقة المتابعة يتم استخدامها بصورة كبيرة في الاقتصاد، فإن المشاكل المترتبة عليها عادة مايتم تجاهلها لسوء الحظ.