

## 4.6 Tự tương quan (Autocorrelation)

Bây giờ chúng ta sẽ xem xét một trường hợp khác khi  $V\{\varepsilon\} = \sigma^2 I$  bị vi phạm, tức là khi hiệp phương sai giữa các sai số không đồng thời bằng 0. Ví dụ, điều này xảy ra khi hai hay nhiều sai số liên tiếp tương quan, trong trường hợp này chúng ta nói rằng xảy ra hiện tượng tự tương quan của sai số (autocorrelation) hay tương quan chuỗi (serial correlation). Như đã thảo luận ở trên, miễn là có thể giả định rằng  $E\{\varepsilon | X\} = 0$  (giả định (A9)), hậu quả của tự tương quan tương tự như phương sai thay đổi (hay phương sai của sai số thay đổi): OLS không thiên lệch, nhưng nó trở nên không hiệu quả và sai số chuẩn của ước lượng các hệ số (SE) bị ước tính sai.

Tự tương quan thường xảy ra chỉ khi sử dụng dữ liệu theo chuỗi thời gian. Để nhấn mạnh điều này, chúng ta sẽ đánh số các quan sát từ  $t = 1, 2, \dots, T$  thay vì từ  $i = 1, 2, \dots, N$ . Sự khác biệt lớn nhất là bây giờ thứ tự của các quan sát trở nên quan trọng và chỉ số phản ánh trật tự tự nhiên. Nhìn chung, sai số  $\varepsilon_t$  ảnh hưởng đến các biến mà những biến này lại ảnh hưởng đến các biến phụ thuộc chưa được đưa vào mô hình. Ảnh hưởng của các biến bị loại trừ này thường là nguyên nhân dẫn đến hiện tượng tự tương quan dương. Nếu các biến loại trừ nói trên được quan sát và được đưa vào mô hình, chúng ta có thể diễn giải kết quả tự tương quan như là một dấu hiệu của một mô hình không chuẩn. Điều này giải thích tại sao các bài kiểm tra sự tự tương quan thường được xem như là kiểm tra sự sai lệch của mô hình. Các dạng hàm mô tả không chính xác, các biến bị bỏ sót và một mô hình thiếu tính động có thể dẫn đến kết luận về sự tự tương quan.

Giả sử bạn đang sử dụng dữ liệu hàng tháng để ước lượng mô hình giải thích nhu cầu về kem. Thông thường, trạng thái thời tiết sẽ là một nhân tố quan trọng ẩn trong sai số  $\varepsilon_t$ . Trong trường hợp này, bạn có thể tìm thấy một mẫu quan sát giống như hình 4.1. Trong mô hình này, chúng tôi tính toán lượng tiêu thụ kem theo thời gian, trong khi các điểm kết nối mô tả các giá trị ước lượng (giá trị thích hợp – fitted values) của một mô hình hồi qui nhằm giải thích lượng tiêu thụ kem thông qua tác động của tổng thu nhập và chỉ số giá.<sup>1</sup> Rõ ràng, các phần dư dương và âm được nhóm cùng với nhau. Trong các phân tích kinh tế vĩ mô, các chu kỳ kinh doanh có thể có những tác động tương tự. Trong hầu hết các ứng dụng kinh tế diễn ra hiện tượng tự tương quan dương, nhưng thỉnh thoảng sẽ có tự tương quan âm: một sai số dương cho một quan sát có thể theo sau bởi một sai số âm trong lần kế tiếp, và ngược lại.

---

<sup>1</sup> Dữ liệu được sử dụng trong tài liệu này được trích từ Hildreth và Lu (1960) và có trong ICECREAM; Xem thêm Phần 4.8.

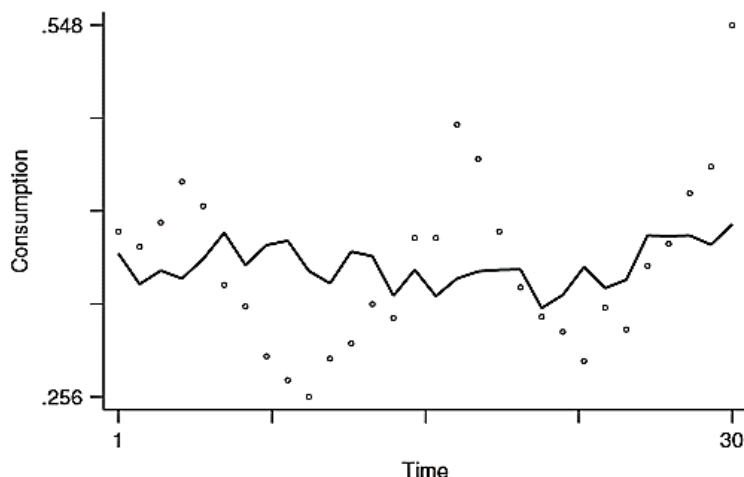


Figure 4.1 Actual and fitted consumption of ice cream, March 1951–July 1953

#### 4.6.1 Tự tương quan bậc nhất (First order autocorrelation)

Có nhiều dạng tự tương quan và mỗi trường hợp sẽ dẫn đến một cấu trúc ma trận hiệp phương sai  $V\{\varepsilon\}$  khác nhau. Hình thức phổ biến nhất được biết đến là quá trình tự hồi qui bậc nhất (first-order autoregressive process). Trong trường hợp này, sai số trong

$$y_t = x_t' \beta + \varepsilon_t \quad (4.47)$$

được giả định phụ thuộc vào

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t, \quad (4.48)$$

trong đó  $v_t$  là sai số với trung bình số học của các sai số bằng 0 và phương sai không đổi  $\sigma_v^2$ , thể hiện không có tương quan chuỗi. Điều này giả định rằng giá trị của sai số trong các quan sát bằng  $\rho$  lần giá trị của nó trong quan sát trước cộng với một thành phần mới  $v_t$ . Thành phần mới này được giả định có trung bình của các sai số bằng 0, phương sai không đổi, và độc lập theo thời gian. Hơn nữa, giả định (A2) ở Chương 2 hàm ý rằng tất cả các biến giải thích độc lập với tất cả các sai số. Các thông số  $\rho$  và  $\sigma_v^2$  thường không biết, và cùng với  $\beta$  chúng ta hy vọng có thể ước lượng được chúng. Lưu ý rằng các thuộc tính thống kê của  $v_t$  cũng giống như các thuộc tính được giả định cho  $\varepsilon_t$  trong trường hợp chuẩn:  $\rho = 0$ ,  $\varepsilon_t = v_t$  và các điều kiện tiêu chuẩn Gauss-Markov (A1) - (A4) ở chương 2 thỏa mãn.

Để tính được ma trận hiệp phương sai của vector sai số  $\varepsilon$ , chúng ta cần có một giả định về sự phân bố của sai số trong giai đoạn đầu tiên,  $\varepsilon_1$ . Thông thường, giả thiết rằng  $\varepsilon_1$  có trung bình bằng 0 và có cùng phương sai với tất cả các sai số  $\varepsilon_t$  khác. Điều này phù hợp với ý tưởng rằng quá trình này đã hoạt động trong một thời gian dài từ quá khứ và  $|\rho| < 1$ . Khi điều kiện  $|\rho| < 1$  được thỏa mãn, chúng ta nói rằng sự tự tương quan bậc nhất có tính dừng (stationary). Một quá trình dừng là khi giá trị trung bình, phương sai và hiệp phương sai của  $\varepsilon_t$  không thay đổi theo thời gian (xem Chương 8 dưới đây). Áp dụng tính dừng, có thể rút ra từ

$$E\{\varepsilon_t\} = \rho E\{\varepsilon_{t-1}\} + E\{v_t\}$$

rằng  $E\{\varepsilon_t\}=0$ . Ngoài ra, từ

$$V\{\varepsilon_t\} = V\{\rho\varepsilon_{t-1} + v_t\} = \rho^2 V\{\varepsilon_{t-1}\} + \sigma_v^2,$$

chúng ta thu được phương sai của  $\varepsilon_t$ , ký hiệu là  $\sigma_\varepsilon^2$ , được cho bởi:

$$\sigma_\varepsilon^2 = V\{\varepsilon_t\} = \frac{\sigma_v^2}{1 - \rho^2}. \quad (4.49)$$

Các yếu tố không nằm trên đường chéo trong ma trận phương sai – hiệp phương sai của  $\varepsilon$ :

$$\text{cov}\{\varepsilon_t, \varepsilon_{t-1}\} = E\{\varepsilon_t \varepsilon_{t-1}\} = \rho E\{\varepsilon_{t-1}^2\} + E\{\varepsilon_{t-1} v_t\} = \rho \frac{\sigma_v^2}{1 - \rho^2}. \quad (4.50)$$

Hiệp phương sai giữa hai sai số của hai thời kỳ khác biệt là:

$$E\{\varepsilon_t \varepsilon_{t-2}\} = \rho E\{\varepsilon_{t-1} \varepsilon_{t-2}\} + E\{\varepsilon_{t-2} v_t\} = \rho^2 \frac{\sigma_v^2}{1 - \rho^2}, \quad (4.51)$$

Và nhìn chung, chúng ta có giá trị không âm của s,

$$E\{\varepsilon_t \varepsilon_{t-s}\} = \rho^s \frac{\sigma_v^2}{1 - \rho^2}. \quad (4.52)$$

Điều này cho thấy với  $0 < |\rho| < 1$  tất cả các phần tử trong  $\varepsilon$  tương quan lẫn nhau với hiệp phương sai giảm dần nếu khoảng thời gian lớn dần (tức là nếu s lớn). Ma trận hiệp phương sai của  $\varepsilon$  là một ma trận đầy đủ (một ma trận không có yếu tố bằng không). Từ ma trận này có thể chuyển đổi thành một ma trận thích hợp hơn, như đã thảo luận trong Phần 4.2. Tuy nhiên, từ các công thức (4.47) và (4.48) có thể thấy được sự chuyển đổi nào là thích hợp. Bởi vì  $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$ , với  $v_t$  thỏa mãn các điều kiện của Gauss-Markov, rõ ràng là một sự chuyển đổi như  $\varepsilon_t - \varepsilon_{t-1}$  sẽ tạo ra những sai số không tự tương quan và có phương sai không thay đổi (hay phương sai của sai số không thay đổi hay phương sai có điều kiện không đổi). Điều này có nghĩa là, tất cả các quan sát nên được chuyển đổi như  $y_t - \rho y_{t-1}$  và  $x_t - \rho x_{t-1}$ . Do đó, mô hình chuyển đổi trở thành

$$y_t - \rho y_{t-1} = (x_t - \rho x_{t-1})' \beta + v_t, \quad t = 2, 3, \dots, T. \quad (4.53)$$

Bởi vì mô hình trong (4.53) thỏa mãn điều kiện của Gauss-Markov, ước lượng OLS sẽ cho kết quả của ước lượng GLS (giả sử  $\rho$  được biết). Tuy nhiên, phát biểu này không phải là hoàn toàn chính xác, vì sự chuyển đổi trong (4.53) không thể áp dụng cho quan sát đầu tiên (vì  $y_0$  và  $x_0$  không được quan sát). Thông tin về quan sát đầu tiên bị thiếu và OLS theo (4.53) chỉ tạo ra một ước tính gần đúng của GLS.<sup>2</sup> Tất nhiên, khi số lượng quan sát lớn, sự thiếu hụt một quan sát thông thường sẽ không ảnh hưởng lớn đến kết quả.

<sup>2</sup> Về mặt kỹ thuật, ma trận chuyển đổi ẩn P được sử dụng ở đây không phải là một ma trận vuông và do đó không khả nghịch.

Quan sát đầu tiên có thể được xử lý bằng cách lưu ý rằng sai số của quan sát đầu tiên,  $\varepsilon_1$ , không tương quan với tất cả các  $v_t$ ,  $t = 2, \dots, T$ . Tuy nhiên, phương sai của  $\varepsilon_1$  (được cho trong (4.49)) lớn hơn nhiều so với phương sai của sai số chuyển đổi ( $v_2, \dots, v_T$ ), đặc biệt khi  $\rho$  gần bằng 1. Để đạt được các sai số không tự tương quan và phương sai không thay đổi trong một mô hình chuyển đổi (bao gồm cả quan sát đầu tiên), quan sát đầu tiên này nên được chuyển đổi bằng cách nhân nó với  $\sqrt{1 - \rho^2}$ . Mô hình chuyển đổi hoàn chỉnh do đó được cho bởi

$$\sqrt{1 - \rho^2}y_1 = \sqrt{1 - \rho^2}x_1'\beta + \sqrt{1 - \rho^2}\varepsilon_1, \quad (4.54)$$

và bởi (4.53) cho các quan sát từ 2 tới T. Dễ dàng nhận ra rằng sai số chuyển đổi trong (4.54) có cùng phương sai với  $v_t$ . OLS áp dụng cho (4.53) và (4.54) tạo ra ước lượng GLS  $\hat{\beta}$ , là ước lượng không thiên lệch tuyến tính tốt nhất (BLUE) cho  $\beta$ .

Nghiên cứu trước (Cochrane và Orcutt, 1949) thường bỏ qua quan sát (chuyển đổi) đầu tiên và ước lượng  $\beta$  từ các quan sát chuyển đổi T - 1 còn lại. Như đã đề cập, điều này chỉ mang lại một ước lượng GLS gần đúng và nó sẽ không được hiệu quả như các ước lượng sử dụng tất cả các quan sát T. Tuy nhiên, nếu T lớn, sự khác biệt giữa hai ước lượng là không đáng kể. Các ước lượng không sử dụng những quan sát chuyển đổi đầu tiên thường được gọi là các ước lượng Cochrane-Orcutt. Tương tự, sự chuyển đổi không bao gồm quan sát đầu tiên được gọi là chuyển đổi Cochrane-Orcutt. Ước lượng sử dụng tất cả các quan sát chuyển đổi thỉnh thoảng được gọi là ước lượng Prais-Winsten (1954).

#### 4.6.2 Không biết giá trị $\rho$

Trong thực tế, rất hiếm khi giá trị của  $\rho$  được biết đến. Trong trường hợp đó, chúng ta sẽ phải ước tính nó. Bắt đầu từ

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t, \quad (4.55)$$

trong đó  $v_t$  thỏa mãn các giả định thông thường, có vẻ như là một điều tự nhiên khi ước tính  $\rho$  từ hồi qui của phần dư OLS  $e_t$  theo  $e_{t-1}$ . Các kết quả ước lượng OLS của  $\rho$  được cho bởi

$$\hat{\rho} = \left( \sum_{t=2}^T e_{t-1}^2 \right)^{-1} \left( \sum_{t=2}^T e_t e_{t-1} \right). \quad (4.56)$$

Mặc dù ước lượng này cho  $\rho$  thường bị thiên lệch, nó là một ước lượng nhất quán cho  $\rho$  trong điều kiện bình thường yếu. Nếu chúng ta sử dụng  $\hat{\rho}$  thay vì  $\rho$  để tính toán ước lượng bình phương tối thiểu tổng quát khả thi (FGLS)  $\hat{\beta}^*$ , thuộc tính BLUE không còn đúng. Trong cùng điều kiện như trước, ước lượng FGLS của  $\hat{\beta}^*$  tiệm cận bằng với ước lượng GLS  $\hat{\beta}$ . Tức là, đối với các mẫu cỡ lớn, chúng ta có thể bỏ qua thực tế là  $\rho$  được ước tính.

Quy trình ước lượng có liên quan là quy trình lặp Cochrane-Orcutt được áp dụng trong nhiều gói phần mềm. Trong quy trình này,  $\rho$  và  $\beta$  được ước lượng theo phương pháp ước lượng đệ quy cho đến khi hội tụ, tức là có ước lượng  $\beta$  với phương pháp FGLS (bằng  $\hat{\beta}^*$ ), các phần

được tính toán lại và  $\rho$  được ước lượng một lần nữa sử dụng các phần dư từ bước FGLS. Với ước lượng mới của  $\rho$ , FGLS được áp dụng lại và đạt được một ước lượng mới của  $\beta$ . Quy trình này tiếp tục cho đến khi hội tụ, tức là cho đến khi cả ước lượng cho  $\rho$  và ước lượng cho  $\beta$  không thay đổi nữa. Có thể mong đợi rằng quy trình này làm tăng tính hiệu quả (nghĩa là giảm phương sai) của ước lượng  $\rho$ . Tuy nhiên, không có gì đảm bảo rằng nó sẽ làm tăng tính hiệu quả của các ước lượng  $\beta$ . Chúng ta biết rằng một cách tiệm cận, sẽ không phải là vấn đề khi ước lượng  $\rho$ , và do đó cũng sẽ không phải là vấn đề (một cách tiệm cận) cách thức làm thế nào để chúng ta có thể ước lượng nó, miễn là nó được tính toán một cách nhất quán. Tuy nhiên, trong các mẫu nhỏ, FGLS lặp đi lặp lại thường thực hiện tốt hơn so với biến thể hai bước của nó.

#### 4.7 Kiểm định sự tự tương quan bậc nhất

Khi  $\rho = 0$ , không có tự tương quan và OLS là BLUE. Nếu  $\rho \neq 0$  suy luận dựa trên ước lượng OLS sẽ bị sai lệch vì sai số chuẩn (SE) dựa trên công thức sai. Do đó, thông thường với dữ liệu chuỗi thời gian cần kiểm tra sự tự tương quan của sai số. Giả sử chúng ta muốn kiểm tra sự tự tương quan bậc nhất biểu thị thông qua  $\rho \neq 0$  trong công thức (4.48). Chúng tôi sẽ trình bày một số kiểm định cho vấn đề tự tương quan trong phần dưới đây. Tập hợp các kiểm định đầu tiên tương đối đơn giản và dựa trên sự xấp xỉ tiệm cận, trong khi kiểm định cuối cùng có phân phối mẫu nhỏ đã biết.

##### 4.7.1 Kiểm định tiệm cận (Asymptotic tests)

Phần dư của OLS từ (4.47) cung cấp thông tin hữu ích về khả năng có mặt của tương quan chuỗi trong các sai số. Một điểm khởi đầu thú vị bằng trực giác là xem xét hồi qui phần dư OLS  $e_t$  theo độ trễ của nó  $e_{t-1}$ . Hồi qui này có thể được thực hiện khi nó bị chặn hoặc không bị chặn (dẫn đến kết quả khác biệt không đáng kể). Các hàm hồi qui bổ sung không chỉ tạo ra một ước tính cho hệ số tự tương quan bậc nhất,  $\hat{\rho}$ , mà thường còn cung cấp một sai số chuẩn cho ước tính này. Trong trường hợp không có các biến phụ thuộc trễ trong (4.47), kiểm định  $t$  tương ứng không có giá trị tiệm cận. Trên thực tế, kết quả của thống kê kiểm định có thể được hiển thị xấp xỉ bằng

$$t \approx \sqrt{T} \hat{\rho}, \quad (4.57)$$

điều này cung cấp một cách khác để tính thống kê kiểm định. Do đó, ở mức ý nghĩa 5%, chúng tôi bác bỏ giả thuyết  $H_0$  rằng không có sự tự tương quan đối với một sự thay thế 2 bên (two-sided alternative) nếu  $|t| > 1,96$ . Nếu giả thuyết  $H_1$  là tự tương quan dương ( $\rho > 0$ ), điều thường được kỳ vọng sẽ xảy ra, giả thuyết  $H_0$  bị loại bỏ ở mức 5% nếu  $t > 1.64$  (so sánh phần 2.5.1).

Một lựa chọn khác là dựa trên  $R^2$  của hàm hồi qui bổ sung (bao gồm cả tính chặn). Nếu chúng ta lấy  $R^2$  của hồi qui này và nhân nó với số hiệu quả của các quan sát  $T-1$  chúng ta có được một kiểm định thống kê mà, theo giả thuyết  $H_0$  có phân phối  $\chi^2$  với một bậc tự do. Rõ ràng, một  $R^2$  gần bằng không trong hồi qui này ngụ ý rằng các phần dư với độ trễ không giải thích các phần dư hiện tại và một cách đơn giản để kiểm tra  $\rho = 0$  là bằng cách tính toán  $(T-1)R^2$ . Kiểm định này là một trường hợp đặc biệt của kiểm định số nhân Lagrange của Breusch

(1978) - Godfrey (1978) (xem chương 6) và nó dễ dàng mở rộng cho sự tự tương quan ở các bậc cao hơn (bằng cách thêm vào độ trễ của các phần dư và điều chỉnh bậc tự do tương ứng).

Nếu mô hình quan tâm bao gồm biến phụ thuộc trễ (hoặc các biến giải thích khác có tương quan với sai số trễ), các kiểm định trên vẫn còn thích hợp với điều kiện các biến hồi qui  $x_t$  được đưa vào hàm hồi qui bổ sung. Điều này cần tính đến khả năng  $x_t$  và  $u_{t-1}$  tương quan và đảm bảo rằng thống kê kiểm định có phân phối xấp xỉ phân phối thích hợp. Khi nghi ngờ rằng sai số trong phương trình quan tâm có phương sai thay đổi, theo đó phương sai của  $\varepsilon_t$  phụ thuộc vào  $x_t$ , các phiên bản kiểm định t của kiểm định sự tự tương quan có thể tạo nên phương sai thay đổi nhất quán (heteroskedasticity-consistent) bằng cách sử dụng sai số chuẩn White (xem Mục 4.3.4) trong hồi qui bổ sung để thực hiện các thống kê kiểm định.

#### 4.7.2 Kiểm định Durbin – Watson

Một kiểm định phổ biến đối với sự tự tương quan bậc nhất là kiểm định Durbin-Watson (Durbin và Watson, 1950), kiểm định này có một phân phối mẫu nhỏ đã biết dưới một tập hợp các điều kiện hạn chế. Hai giả định quan trọng cho thử nghiệm này là chúng ta có thể xem  $x_t$  là tất định và  $x_t$  có chứa thành phần bị chặn. Giả định đầu tiên là quan trọng bởi vì nó yêu cầu tất cả các sai số độc lập với tất cả các biến giải thích (giả định (A2)). Quan trọng nhất là, điều này loại trừ việc bao gồm biến phụ thuộc trễ trong mô hình.

Thống kê kiểm định Durbin-Watson được cho bởi

$$dw = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}, \quad (4.58)$$

trong đó  $e_t$  là phần dư của OLS (chú ý các chỉ số khác nhau cho việc lấy tổng). Đại số đơn giản chỉ ra rằng

$$dw \approx 2 - 2\hat{\rho}, \quad (4.59)$$

dấu xấp xỉ để chỉ sự khác biệt nhỏ trong các quan sát mà phép tổng được thực hiện. Do đó, giá trị của  $d_w$  gần bằng 2 chỉ ra rằng hệ số tự tương quan bậc nhất  $\rho$  gần bằng không. Nếu  $d_w$  'nhỏ hơn nhiều' so với 2, điều này ám chỉ một sự tương quan dương ( $\rho > 0$ ); nếu  $d_w$  lớn hơn nhiều so với 2 thì  $\rho < 0$ . Ngay cả với  $H_0: \rho = 0$ , phân phối  $d_w$  phụ thuộc không chỉ vào kích cỡ mẫu  $T$  và số lượng các biến  $K$  trong  $x_t$ , mà còn phụ thuộc vào giá trị thực của  $x_t$ s. Do đó, các giá trị quan trọng không thể được lập thành bảng để sử dụng chung. May mắn là, có thể tính các giới hạn trên và giới hạn dưới cho các giá trị quan trọng của  $d_w$  mà các giá trị này chỉ phụ thuộc vào kích cỡ mẫu  $T$  và số lượng biến  $K$  trong  $x_t$ . Các giá trị này,  $d_L$  và  $d_U$ , được Durbin và Watson (1950) và Savin and White (1977) lập thành bảng, và một phần trong số đó được trình bày trong Bảng 4.8. Giới hạn tin cậy (hay giá trị tới hạn)  $d_{crit}$  nằm giữa các giới hạn được lập bảng, hay  $d_L < d_{crit} < d_U$ . Theo giả thiết  $H_0$  chúng ta có (mức ý nghĩa 5%)

$$P\{dw < d_L\} \leq P\{dw < d_{crit}\} = 0.05 \leq P\{dw < d_U\}.$$

**Table 4.8** Lower and upper bounds for 5% critical values of the Durbin–Watson test (Savin and White, 1977)

| Number of observations | Number of regressors (incl. intercept) |       |         |       |         |       |         |       |
|------------------------|----------------------------------------|-------|---------|-------|---------|-------|---------|-------|
|                        | $K = 3$                                |       | $K = 5$ |       | $K = 7$ |       | $K = 9$ |       |
|                        | $d_L$                                  | $d_U$ | $d_L$   | $d_U$ | $d_L$   | $d_U$ | $d_L$   | $d_U$ |
| $T = 25$               | 1.206                                  | 1.550 | 1.038   | 1.767 | 0.868   | 2.012 | 0.702   | 2.280 |
| $T = 50$               | 1.462                                  | 1.628 | 1.378   | 1.721 | 1.291   | 1.822 | 1.201   | 1.930 |
| $T = 75$               | 1.571                                  | 1.680 | 1.515   | 1.739 | 1.458   | 1.801 | 1.399   | 1.867 |
| $T = 100$              | 1.634                                  | 1.715 | 1.592   | 1.758 | 1.550   | 1.803 | 1.506   | 1.850 |
| $T = 200$              | 1.748                                  | 1.789 | 1.728   | 1.810 | 1.707   | 1.831 | 1.686   | 1.852 |

Đối với kiểm định một bên (one-sided test) kiểm tra sự tự tương quan dương ( $\rho > 0$ ), có ba khả năng:

- $d_w$  nhỏ hơn  $d_L$ . Trong trường hợp này, nó chắc chắn thấp hơn giới hạn tin cậy  $d_{crit}$ , vì vậy bạn sẽ bác bỏ  $H_0$ ;
- $d_w$  lớn hơn  $d_U$ . Trong trường hợp này, nó chắc chắn lớn hơn  $d_{crit}$  và bạn sẽ không bác bỏ  $H_0$ ;
- $d_w$  nằm giữa  $d_L$  và  $d_U$ . Trong trường hợp này nó có thể lớn hơn hoặc nhỏ hơn giới hạn tin cậy  $d_{crit}$ . Bởi vì bạn không thể nói trường hợp nào xảy ra, bạn không thể chấp nhận hoặc bác bỏ  $H_0$ . Đây gọi là 'khu vực không thể quyết định'.

Kích thước mẫu càng lớn thì khu vực không thể kết luận càng nhỏ. Khi  $K = 5$  và  $T = 25$  chúng ta có  $d_{L, 5\%} = 1,038$  và  $d_{U, 5\%} = 1,767$ ; khi  $T = 100$  những con số này là 1,592 và 1,758.

Sự tồn tại của một vùng tổng thể và yêu cầu rằng các điều kiện Gauss-Markov, bao gồm cả tính tiêu chuẩn của sai số, phải được thỏa mãn là những hạn chế quan trọng của kiểm định Durbin-Watson. Tuy nhiên, bởi vì nó thường được cung cấp bởi hầu hết các gói hồi qui, nó thường sẽ nhanh chóng đưa ra dấu hiệu cho khả năng hiện diện của sự tự tương quan. Các giá trị nhỏ hơn đáng kể so với 2 là một dấu hiệu của sự tự tương quan dương (vì chúng tương ứng với  $\hat{\rho} > 0$ ). Lưu ý rằng các phép kiểm tra tiệm cận là gần đúng, thậm chí khi không có các sai số thông thường, và có thể được mở rộng cho các biến phụ thuộc trễ trong  $x_t$ .

Trong trường hợp ít phổ biến hơn nếu giả thuyết  $H_1$  là sự hiện diện của sự tự tương quan âm ( $\rho < 0$ ), giới hạn tin cậy là giữa  $4 - d_U$  và  $4 - d_L$ , để khỏi yêu cầu thêm các bảng khác.

#### 4.8 Minh họa: Nhu cầu về kem

Minh họa thực nghiệm này dựa trên một trong những bài viết về sự tự tương quan, Hildreth và Lu (1960). Dữ liệu được sử dụng trong nghiên cứu này là dữ liệu chuỗi thời gian với 30 quan sát trong bốn tuần từ ngày 18 tháng 3 năm 1951 đến ngày 11 tháng 7 năm 1953 về các biến sau:<sup>3</sup>

cons: lượng tiêu thụ kem trên đầu người (theo pint);

income: thu nhập trung bình của các gia đình hàng tuần (đô la Mỹ);

<sup>3</sup> Dữ liệu có sẵn trên ICECREAM.

price: giá kem (theo pint);

temp: nhiệt độ trung bình (theo Fahrenheit).

Một đồ thị minh họa cho dữ liệu được thể hiện trong hình 4.2 với mô hình chuỗi thời gian của tiêu dùng, giá cả và nhiệt độ (chia cho 100). Biểu đồ gợi ý rằng nhiệt độ là yếu tố quyết định đối với lượng tiêu thụ kem, điều này hỗ trợ cho kỳ vọng của chúng tôi.

Mô hình được sử dụng để giải thích lượng tiêu thụ kem là mô hình hồi qui tuyến tính với thu nhập (income), giá cả (price) và nhiệt độ (temp) là các biến giải thích. Kết quả hồi qui OLS đầu tiên được trình bày trong Bảng 4.9. Mặc dù các hệ số ước lượng có dấu như dự kiến, thống kê Durbin-Watson được tính là 1.0212. Đối với kiểm định Durbin-Watson một bên với  $H_0: \rho = 0$ , và  $H_1$  là sự tự tương quan dương, chúng tôi có ở ý nghĩa mức 5% ( $\alpha = 0,05$ ) rằng  $d_L = 1,21$  ( $T = 30$ ,  $K = 4$ ) và  $d_U = 1,65$ . Giá trị 1.02 rõ ràng hàm ý rằng giả thuyết  $H_0$  nên bị bác bỏ so với giả thiết  $H_1$  rằng có sự tự tương quan dương. Khi chúng tôi vẽ các giá trị quan sát của cons và giá trị dự đoán theo mô hình, như trong hình 4.3, chúng tôi thấy rằng các giá trị dương (âm) của sai số có nhiều khả năng sẽ được theo sau bởi các giá trị dương (âm). Rõ ràng, sự bao gồm nhiệt độ trong mô hình không đủ để nắm bắt được sự biến động theo mùa trong việc tiêu thụ kem.

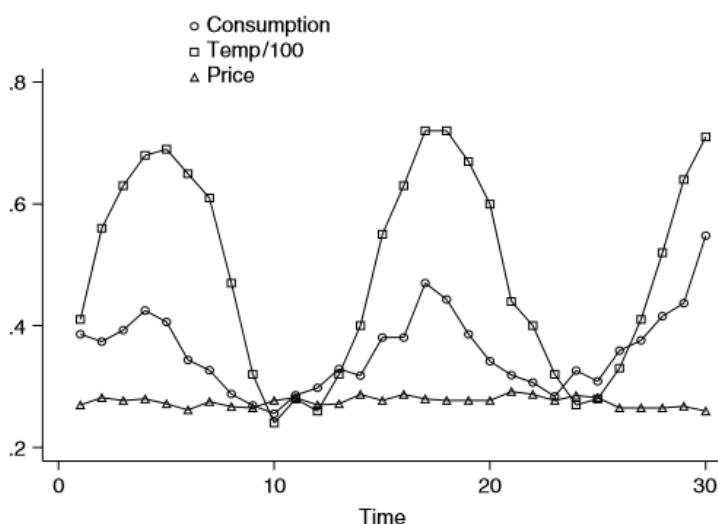


Figure 4.2 Ice cream consumption, price and temperature/100

Table 4.9 OLS results

| Dependent variable: <i>cons</i>                               |          |                |                 |
|---------------------------------------------------------------|----------|----------------|-----------------|
| Variable                                                      | Estimate | Standard error | <i>t</i> -ratio |
| constant                                                      | 0.197    | 0.270          | 0.730           |
| <i>price</i>                                                  | -1.044   | 0.834          | -1.252          |
| <i>income</i>                                                 | 0.00331  | 0.00117        | 2.824           |
| <i>temp</i>                                                   | 0.00345  | 0.00045        | 7.762           |
| $s = 0.0368$ $R^2 = 0.7190$ $\bar{R}^2 = 0.6866$ $F = 22.175$ |          |                |                 |
| $dw = 1.0212$                                                 |          |                |                 |



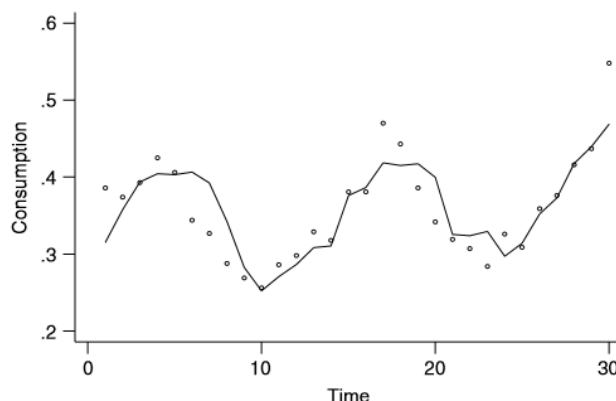


Figure 4.3 Actual and fitted values (connected) ice cream consumption

Hệ số tự tương quan bậc nhất trong

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t$$

được ước lượng một cách dễ dàng bằng cách giữ lại phần dư từ hồi qui trước và chạy một hồi qui bình phương tối thiểu của  $\varepsilon_t$  theo  $\varepsilon_{t-1}$  (không có hằng số).<sup>4</sup> Điều này cho một ước tính  $\hat{\rho} = 0.401$  với  $R^2$  là 0.149. Kiểm định tiệm cận đối với  $H_0: \rho = 0$  đối với tự tương quan bậc nhất dựa trên  $\sqrt{T}\hat{\rho} = 2.19$ . Giá trị này lớn hơn giới hạn tin cậy 5% của phân phối chuẩn là 1.96, vì vậy chúng ta bác bỏ giả thiết  $H_0$  không có tương quan chuỗi. Các thử nghiệm Breusch-Godfrey tạo ra một kiểm định thống kê của  $(T-1)R^2 = 4.32$ , vượt quá giá trị tới hạn 5% là 3.84 của p Chi-squared với một bậc tự do.

Những sự bác bỏ này hàm ý rằng OLS không còn là ước lượng tuyến tính không thiên lệch tốt nhất cho  $\beta$  và, quan trọng nhất, là các sai số chuẩn thường được tính toán không chính xác. Có thể đưa ra những tuyên bố chính xác về độ co giãn theo giá cả của kem nếu chúng ta chọn một phương pháp ước lượng hiệu quả hơn, như GLS. Quá trình Cochrane-Orcutt lặp cho kết quả được trình bày trong Bảng 4.10. Lưu ý rằng các kết quả FGLS xác nhận kết quả trước đó của chúng tôi, điều này cho thấy rằng thu nhập và nhiệt độ là những yếu tố quyết định trong hàm tiêu thụ. Cần nhấn mạnh rằng các số liệu thống kê trong Bảng 4.10 có kèm theo dấu sao (\*) tương ứng với mô hình chuyển đổi và không trực tiếp so sánh với các mô hình tương đương trong Bảng 4.9 là các mô hình không chuyển đổi. Điều này cũng giống như trong thống kê của Durbin-Watson, điều không còn phù hợp trong Bảng 4.10.

<sup>4</sup> Không cần bao gồm một hằng số bởi vì trung bình phần dư trong ước lượng OLS bằng 0.

**Table 4.10** EGLS (iterative Cochrane–Orcutt) results

| Dependent variable: <i>cons</i>                                     |          |                |                 |
|---------------------------------------------------------------------|----------|----------------|-----------------|
| Variable                                                            | Estimate | Standard error | <i>t</i> -ratio |
| constant                                                            | 0.157    | 0.300          | 0.524           |
| <i>price</i>                                                        | −0.892   | 0.830          | −1.076          |
| <i>income</i>                                                       | 0.00320  | 0.00159        | 2.005           |
| <i>temp</i>                                                         | 0.00356  | 0.00061        | 5.800           |
| $\hat{\rho}$                                                        | 0.401    | 0.2079         | 1.927           |
| $s = 0.0326^*$ $R^2 = 0.7961^*$ $\bar{R}^2 = 0.7621^*$ $F = 23.419$ |          |                |                 |
| $dw = 1.5486^*$                                                     |          |                |                 |

**Table 4.11** OLS results extended specification

| Dependent variable: <i>cons</i>                               |          |                |                 |
|---------------------------------------------------------------|----------|----------------|-----------------|
| Variable                                                      | Estimate | Standard error | <i>t</i> -ratio |
| constant                                                      | 0.189    | 0.232          | 0.816           |
| <i>price</i>                                                  | −0.838   | 0.688          | −1.218          |
| <i>income</i>                                                 | 0.00287  | 0.00105        | 2.722           |
| <i>temp</i>                                                   | 0.00533  | 0.00067        | 7.953           |
| <i>temp</i> <sub><i>t</i>−1</sub>                             | −0.00220 | 0.00073        | −3.016          |
| $s = 0.0299$ $R^2 = 0.8285$ $\bar{R}^2 = 0.7999$ $F = 28.979$ |          |                |                 |
| $dw = 1.5822$                                                 |          |                |                 |

Như đã đề cập ở trên, việc tìm ra sự tự tương quan có thể là một dấu hiệu cho thấy có sai sót trong mô hình, ví dụ như dạng hàm mô tả không chính xác hoặc một mô hình với đặc trưng thiếu tính động. Một cách để có thể loại bỏ các vấn đề của sự tự tương quan là thay đổi đặc tính của mô hình. Có vẻ là tự nhiên khi xem xét việc bao gồm một hoặc nhiều biến trễ trong mô hình. Đặc biệt, chúng ta sẽ bao gồm biến trễ về nhiệt độ  $temp_{t-1}$  trong mô hình. Ước lượng OLS của mô hình mở rộng này cho kết quả như trong Bảng 4.11.

So với Bảng 4.9, kiểm định thống kê Durbin-Watson đã tăng lên 1,58, nằm ở vùng không thể quyết định ( $\alpha = 0,05$ ) với (1,14, 1,74). Do giá trị gần tương đương với giới hạn trên, chúng ta có thể chọn không bác bỏ giả thiết  $H_0$  rằng không có sự tự tương quan. Rõ ràng độ trễ của biến nhiệt độ có ảnh hưởng tiêu cực đến việc tiêu thụ kem, trong khi nhiệt độ hiện tại lại có ảnh hưởng tích cực. Điều này có thể ám chỉ rằng có một sự gia tăng nhu cầu về kem khi nhiệt độ tăng, đây là sự tiêu thụ không hoàn toàn và làm giảm chi phí giai đoạn sau đó.<sup>5</sup>

## 4.9 Các mô hình tự tương quan khác

### 4.9.1 Tự tương quan bậc cao hơn

Sai số tự tương quan bậc nhất khá phổ biến trong các mô hình chuỗi thời gian trong kinh tế vĩ mô và trong hầu hết các trường hợp, việc cho phép tự tương quan sẽ loại bỏ vấn đề. Tuy nhiên, ví dụ như khi chúng ta có dữ liệu hàng quý hoặc hàng tháng, có thể có ảnh hưởng chu kỳ (hàng quý hoặc hàng tháng) gây ra sai số trong những giai đoạn giống nhau nhưng trong

<sup>5</sup> Biến *cons* được đo lường bằng chi tiêu cho kem, không phải là lượng tiêu dùng thực tế.

những năm khác nhau có mối tương quan. Ví dụ: chúng ta có thể có (trong trường hợp dữ liệu hàng quý)

$$\varepsilon_t = \gamma \varepsilon_{t-4} + v_t, \quad (4.60)$$

hoặc khái quát hơn là

$$\varepsilon_t = \gamma_1 \varepsilon_{t-1} + \gamma_2 \varepsilon_{t-2} + \gamma_3 \varepsilon_{t-3} + \gamma_4 \varepsilon_{t-4} + v_t, \quad (4.61)$$

được gọi là tự tương quan bậc bốn. Về bản chất, đây là một khái quát đơn giản của quá trình bậc nhất và ước lượng FGLS tương tự như vậy. Miễn là các biến giải thích không tương quan với tất cả các sai số, FGLS dựa trên bước ước lượng OLS đầu tiên của (4.60) hoặc (4.61), trong đó các sai số được thay thế bởi các phần dư với bình phương tối thiểu  $e_t$ . Việc chuyển đổi thích hợp để đạt được ước lượng FGLS cho  $\beta$  sẽ rõ ràng từ (4.60) hoặc (4.61). Lưu ý rằng bốn quan sát đầu tiên sẽ bị mất trong quá trình chuyển đổi.

#### 4.9.2 Sai số trung bình trượt (Sai số trung bình dịch chuyển)

Như đã thảo luận, một đặc trưng tự hồi qui của sai số, như trong (4.48), (4.60) hoặc (4.61) ám chỉ rằng tất cả sai số tương quan lẫn nhau, mặc dù sự tương quan giữa các sai số cách xa nhau về thời gian là không đáng kể. Trong một số trường hợp, lý thuyết (kinh tế) gợi ý một hình thức khác của tự tương quan, trong đó chỉ có những sai số nhất định là tương quan, trong khi tất cả những sai số khác không tương quan với nhau. Điều này có thể được mô hình hóa bằng một quá trình gọi là sai số trung bình trượt. Sai số trung bình trượt thường phát sinh khi khoảng lấy mẫu (ví dụ: một tháng) nhỏ hơn khoảng thời gian mà các biến được xác định. Xem xét vấn đề ước lượng một phương trình giải thích giá trị của một số công cụ tài chính như tín phiếu kho bạc 90 ngày hoặc hợp đồng kỳ hạn 3 tháng về ngoại hối. Nếu sử dụng dữ liệu theo tháng, thì bất kỳ sự đổi mới nào xảy ra trong tháng  $t$  sẽ ảnh hưởng đến giá trị của các công cụ tới hạn trong tháng  $t$ ,  $t+1$  và  $t+2$  nhưng sẽ không ảnh hưởng đến giá trị của các công cụ tới hạn sau đó, bởi vì nó vẫn chưa được ban hành. Điều này cho thấy sự tương quan giữa các sai số cách nhau một và hai tháng, nhưng không tương quan với các sai số xa hơn nữa.

Một ví dụ khác là lời giải thích về sự thay đổi giá cả hàng năm (lạm phát), được quan sát 6 tháng một lần. Giả sử chúng ta quan sát sự thay đổi trong giá tiêu dùng so với mức một năm trước đó, vào ngày 1 tháng Một và ngày 1 Tháng Bảy. Cũng giả sử rằng các biến số cơ sở (ví dụ cung tiền) bao gồm trong  $x_t$  được quan sát nửa năm. Nếu mô hình 'đúng' được cho bởi

$$y_t = x_t' \beta + v_t, \quad t = 1, 2, \dots, T \text{ (half-yearly)}, \quad (4.62)$$

với  $y_t$  là sự thay đổi giá cả trong nửa năm và sai số  $v_t$  đáp ứng các điều kiện Gauss-Markov, điều này vẫn đúng cho sự thay đổi theo năm,  $y_t^* = y_t + y_{t-1}$  mà

$$y_t^* = (x_t + x_{t-1})' \beta + v_t + v_{t-1}, \quad t = 1, 2, \dots, T, \quad (4.63)$$

hoặc

$$y_t^* = x_t^* \beta + \varepsilon_t, \quad t = 1, 2, \dots, T, \quad (4.64)$$

với  $\varepsilon_t = v_t + v_{t-1}$  và  $x_t^* = x_t + x_{t-1}$ . Nếu chúng ta giả định rằng  $v_t$  có phương sai  $\sigma_v^2$ , các thuộc tính của sai số trong công thức (4.64) là như sau

$$\begin{aligned} E\{\varepsilon_t\} &= E\{v_t\} + E\{v_{t-1}\} = 0 \\ V\{\varepsilon_t\} &= V\{v_t + v_{t-1}\} = 2\sigma_v^2 \\ \text{cov}\{\varepsilon_t, \varepsilon_{t-1}\} &= \text{cov}\{v_t + v_{t-1}, v_{t-1} + v_{t-2}\} = \sigma_v^2 \\ \text{cov}\{\varepsilon_t, \varepsilon_{t-s}\} &= \text{cov}\{v_t + v_{t-1}, v_{t-s} + v_{t-1-s}\} = 0, \quad s = 2, 3, \dots \end{aligned}$$

Do đó, ma trận hiệp phương sai của vector sai số chứa một số lượng lớn các số không. Trên đường chéo chúng ta có  $2\sigma_v^2$  (phương sai) và ngay dưới và trên đường chéo chúng ta có  $\sigma_v^2$  (tự hiệp phương sai bậc nhất - autocovariance), trong khi tất cả các hiệp phương sai khác bằng không. Chúng ta gọi đây là một quy trình trung bình trượt bậc nhất (cho  $\varepsilon_t$ ). Trên thực tế, đây là một phiên bản bị giới hạn bởi vì hệ số tương quan giữa  $\varepsilon_t$  và  $\varepsilon_{t-1}$  theo tiên nghiệm được xác định ở mức 0,5. Một quy trình trung bình trượt bậc nhất nhìn chung sẽ được cụ thể hóa như sau:

$$\varepsilon_t = v_t + \alpha v_{t-1}$$

cho một số  $\alpha$ ,  $|\alpha| < 1$ ; xem thảo luận trong Chương 8 về các mô hình chuỗi thời gian.

Nhìn chung, thường khó ước lượng các mô hình hồi qui với sai số trung bình trượt hơn so với mô hình với quy trình tự hồi qui. Điều này là do sự chuyển đổi tạo ra 'sai số Gauss-Markov' khá phức tạp. Một số gói phần mềm có các thủ tục chuyên biệt có sẵn, nhưng nếu thiếu phần mềm thích hợp, việc ước tính có thể khá khó khăn. Một giải pháp có thể thực hiện được là áp dụng phương pháp bình phương tối thiểu thông thường trong khi điều chỉnh sai số chuẩn cho sự có mặt của tự tương quan (của bất cứ đặc tính nào) trong  $\varepsilon_t$ . Điều này sẽ được thảo luận trong phần tiếp theo. Một ví dụ thực tế liên quan đến sai số trung bình dịch chuyển được nêu trong Phần 4.11.

#### 4.10 Cần làm gì khi bạn phát hiện sự tự tương quan

Trong nhiều trường hợp, việc tìm ra sự tự tương quan là một dấu hiệu cho thấy mô hình không chuẩn. Nếu đúng như vậy, cách tự nhiên nhất không phải là thay đổi ước lượng của bạn (từ OLS sang FGLS) mà là thay đổi mô hình của bạn. Thông thường, ba loại mô hình không chuẩn (có liên quan với nhau) có thể dẫn đến việc xuất hiện tự tương quan trong phần dư của ước lượng OLS: việc thiếu tính động trong mô hình, bỏ sót biến và dạng hàm mô tả không chính xác.

Nếu chúng ta bỏ qua trường hợp sai số độc lập với tất cả các biến giải thích, có một lý do khác tại sao GLS hoặc FGLS có thể không phù hợp. Cụ thể là, có thể giả định rằng ước lượng GLS là không nhất quán vì mô hình chuyển đổi không đáp ứng các yêu cầu tối thiểu để ước lượng OLS nhất quán. Tình huống này có thể nảy sinh ngay cả khi OLS áp dụng cho phương trình ban đầu là nhất quán. Phần 4.11 cung cấp một ví dụ thực nghiệm về vấn đề này.

### 4.10.1 Việc xác định sai

Chúng ta hãy bắt đầu với sự sai sót về dạng hàm mô tả. Giả sử rằng mối quan hệ tuyến tính thực sự là giữa  $y_t$  và  $\log x_t$

$$y_t = \beta_1 + \beta_2 \log x_t + \varepsilon_t$$

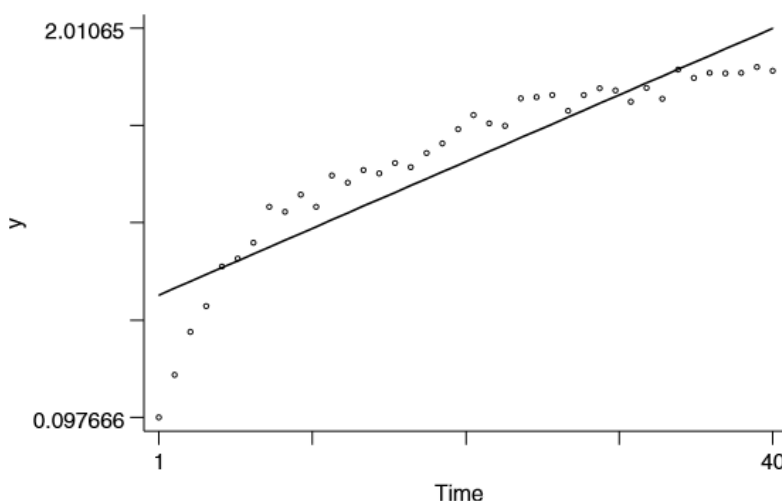


Figure 4.4 Actual and fitted values when true model is  $y_t = 0.5 \log t + \varepsilon_t$

và giả sử, với mục đích minh họa,  $x_t$  tăng theo  $t$ . Tuy nhiên, nếu chúng ta ước lượng một mô hình tuyến tính giải thích  $y_t$  theo  $x_t$  chúng ta có thể tìm thấy một tình huống như mô tả trong hình 4.4. Trong mô hình này, dựa trên dữ liệu mô phỏng với  $x_t = t$  và  $y_t = 0.5 \log x_t$  cộng với một sai số nhỏ, các giá trị ước lượng của mô hình tuyến tính được nối với nhau trong khi các giá trị thực tế thì không. Rõ ràng là, phần dư có cùng dấu được nhóm với nhau. Các thống kê Durbin-Watson tương ứng với ví dụ này bằng 0,193. Giải pháp trong trường hợp này không phải là ước lượng lại mô hình tuyến tính sử dụng bình phương tối thiểu tổng quát khả thi mà là thay đổi dạng hàm mô tả thành  $\log x_t$  chứ không phải là  $x_t$ .

Như đã thảo luận ở trên, sự thiếu sót của một biến giải thích có liên quan cũng có thể làm xuất hiện sự tự tương quan. Ví dụ, trong Phần 4.8 chúng ta thấy rằng việc loại trừ các biến phản ánh sự biến đổi theo mùa của việc tiêu thụ kem đã dẫn đến trường hợp như vậy. Tương tự, việc thiếu tính động trong mô hình cũng có thể dẫn đến sự tự tương quan. Trong những trường hợp như vậy, chúng ta phải quyết định xem liệu mô hình cần quan tâm là tĩnh hoặc động. Để minh họa điều này, bắt đầu từ mô hình (tĩnh)

$$y_t = x_t' \beta + \varepsilon_t \quad (4.65)$$

với tự tương quan bậc nhất  $\varepsilon_t = \rho \varepsilon_{t-1} + v_t$ . Chúng ta có thể giải thích mô hình trên như mô tả  $E\{y_t | x_t\} = x_t' \beta$ . Tuy nhiên, chúng ta cũng có thể quan tâm đến dự báo dựa trên các giá trị  $x_t$  hiện tại cũng như các quan sát với độ trễ  $x_{t-1}$  và  $y_{t-1}$ , nghĩa là  $E\{y_t | x_t, x_{t-1}, y_{t-1}\}$ . Đối với mô hình trên, chúng ta có được

$$E\{y_t|x_t, x_{t-1}, y_{t-1}\} = x_t'\beta + \rho(y_{t-1} - x_{t-1}'\beta) \quad (4.66)$$

và chúng ta có thể viết lại một mô hình động là

$$y_t = x_t'\beta + \rho y_{t-1} - \rho x_{t-1}'\beta + v_t, \quad (4.67)$$

sai số không thể hiện bất kỳ sự tương quan nào. Mô hình (4.67) cho thấy rằng việc bao gồm một biến phụ thuộc trễ và các biến ngoại sinh trễ giúp tránh khỏi hiện tượng tự tương quan. Ngược lại, chúng ta có thể tìm thấy sự tự tương quan trong (4.65) nếu đặc tính động tương tự như (4.67) nhưng, ví dụ, nó chỉ bao gồm  $y_{t-1}$  hoặc một số phần tử của  $x_{t-1}$ . Trong những trường hợp như vậy, việc bao gồm các biến “bị bỏ sót” này sẽ giải quyết vấn đề tự tương quan.

Mô hình tĩnh (4.65) với tự tương quan bậc nhất cung cấp cho chúng ta  $E\{y_t|x_t\}$  cũng như dự báo động  $E\{y_t|x_t, x_{t-1}, y_{t-1}\}$  và có thể hạn chế hơn so với một mô hình động đầy đủ với một số biến số trễ được bao gồm (với các hệ số không giới hạn). Đó là vấn đề lựa chọn cho dù chúng ta quan tâm đến  $E\{y_t|x_t\}$  hay  $E\{y_t|x_t, x_{t-1}, y_{t-1}\}$  hoặc cả hai. Ví dụ, giải thích mức lương của một người dựa trên mức lương của anh ấy trong năm trước có thể khá dễ, nhưng có thể không cung cấp câu trả lời cho những câu hỏi mà chúng ta quan tâm. Tuy nhiên, trong nhiều ứng dụng, việc bao gồm một biến phụ thuộc trễ trong mô hình sẽ loại bỏ vấn đề tự tương quan. Tuy nhiên, cần nhấn mạnh rằng thống kê kiểm định Durbin-Watson không thích hợp trong một mô hình với sự có mặt của biến phụ thuộc trễ. Trong phần 5.2.1 sự chú ý được tập trung vào các mô hình có cả tự tương quan và biến phụ thuộc trễ.

#### 4.10.2 Sai số chuẩn phương sai thay đổi và tự tương quan nhất quán trong ước lượng OLS (sai số chuẩn HAC)

Xem xét lại mô hình căn bản của chúng ta

$$y_t = x_t'\beta + \varepsilon_t, \quad (4.68)$$

với  $\varepsilon_t$  tự tương quan. Ví dụ, nếu đây là mô hình mà chúng ta quan tâm, bởi vì chúng ta muốn biết kỳ vọng có điều kiện của  $y_t$  với  $x_t$  xác định, chúng ta có thể lựa chọn áp dụng phương pháp GLS hoặc áp dụng phương pháp bình phương tối thiểu thông thường trong khi điều chỉnh các sai số chuẩn của nó. Cách tiếp cận cuối cùng này đặc biệt hữu ích khi tương quan giữa  $\varepsilon_t$  và  $\varepsilon_{t-s}$  có thể được lập luận là (hầu như) bằng không sau một số độ trễ có độ dài là  $H$  và/hoặc khi điều kiện về tính nhất quán của ước lượng GLS bị vi phạm.

Nếu  $E\{x_t\varepsilon_t\} = 0$  và  $E\{\varepsilon_t\varepsilon_{t-s}\} = 0$  với  $s = H, H + 1, \dots$ , ước lượng OLS là nhất quán và ma trận hiệp phương sai của nó có thể được ước tính bởi công thức

$$\hat{V}^*\{b\} = \left(\sum_{t=1}^T x_t x_t'\right)^{-1} T S^* \left(\sum_{t=1}^T x_t x_t'\right)^{-1}, \quad (4.69)$$

với

$$S^* = \frac{1}{T} \sum_{t=1}^T e_t^2 x_t x_t' + \frac{1}{T} \sum_{j=1}^{H-1} w_j \sum_{s=j+1}^T e_s e_{s-j} (x_s x_{s-j}' + x_{s-j} x_s'). \quad (4.70)$$

Lưu ý rằng chúng ta có được ma trận hiệp phương sai White, như đã thảo luận trong phần 4.3.4, nếu  $w_j = 0$ , đề (4.69) là một công thức khái quát. Trong trường hợp tiêu chuẩn  $w_j = 1$ , nhưng điều này có thể dẫn đến ma trận hiệp phương sai được ước lượng với số lượng mẫu giới hạn không phải là ma trận xác định dương. Để ngăn chặn điều này, cách phổ biến là sử dụng trọng số Bartlett, theo gợi ý của Newey và West (1987). Các trọng số này giảm tuyến tính với  $j$ ,  $w_j = 1 - j/H$ . Việc sử dụng một tập hợp các trọng số phù hợp với ý tưởng rằng tác động của tự tương quan của trật tự  $j$  giảm theo  $|j|$ . Sai số chuẩn được tính theo công thức (4.69) được gọi là sai số chuẩn phương sai thay đổi và tự tương quan nhất quán (HAC) hoặc đơn giản là sai số chuẩn Newey-West. Thành thạo các sai số chuẩn HAC cũng được sử dụng khi sự tự tương quan không giới hạn bởi độ trễ  $H$ , ví dụ với một cấu trúc tự hồi qui. Về mặt lý thuyết, điều này có thể được điều chỉnh bằng cách áp dụng một đối số tiệm cận mà  $H$  tăng theo  $T$  khi  $T$  tiến đến vô cực (nhưng không nhanh bằng  $T$ ). Theo kinh nghiệm, điều này có thể không hoạt động hiệu quả khi mẫu nhỏ.

Để đạt được một số trực giác cho biểu thức (4.70), cần lưu ý rằng  $S^*$  là một ước lượng cho ma trận hiệp phương sai tiệm cận của trung bình mẫu  $(1/T) \sum_{t=1}^T x_t \varepsilon_t$  (so sánh (2.33) trong Chương 2). Giả sử rằng  $\varepsilon_t$  có thể quan sát được, một người có thể nghĩ đến việc ước lượng ma trận hiệp phương sai này như sau

$$\frac{1}{T} \sum_{s,t} \varepsilon_t \varepsilon_s x_t x_s',$$

với tổng là tất cả các yếu tố có liên quan (đối xứng theo  $s$  và  $t$ ). Ước lượng này không thật sự nhất quán, bởi vì, ví dụ, hiệp phương sai giữa  $x_t \varepsilon_t$  và  $x_{t+H} \varepsilon_{t+H}$  được ước lượng từ một điểm dữ liệu. Điều này giải thích tại sao chúng ta cần phải hạn chế cấu trúc tự tương quan. Với việc không có tự tương quan tại độ dài của độ trễ là  $H$  hoặc lớn hơn, tổng chỉ hơn  $|s-t| \leq H-1$  và ước tính trên trở nên nhất quán.

Trọng số Bartlett đảm bảo rằng các ước tính  $S^*$  là xác định dương trong mỗi mẫu. Điều này có thể được hiểu bằng cách nhìn vào ma trận hiệp phương sai của tổng trong một thời gian ngắn  $\sum_{j=0}^{H-1} x_{t-j} \varepsilon_{t-j}$ , được cho bởi

$$\begin{aligned} V \left\{ \sum_{j=0}^{H-1} x_{t-j} \varepsilon_{t-j} \right\} &= HE \{ \varepsilon_t^2 x_t x_t' \} \\ &\quad + (H-1)[E \{ \varepsilon_t \varepsilon_{t-1} x_t x_{t-1}' \} + E \{ \varepsilon_{t-1} \varepsilon_t x_{t-1} x_t' \}] + \dots \\ &\quad + [E \{ \varepsilon_t \varepsilon_{t-H+1} x_t x_{t-H+1}' \} + E \{ \varepsilon_{t-H+1} \varepsilon_t x_{t-H+1} x_t' \}] \\ &= H \sum_{j=0}^{H-1} (1 - j/H) [E \{ \varepsilon_t \varepsilon_{t-j} x_t x_{t-j}' \} + E \{ \varepsilon_{t-j} \varepsilon_t x_{t-j} x_t' \}], \end{aligned}$$

theo cách xây dựng, đây là ma trận xác định dương. Chia cho  $H$ , và thay thế các toán tử mong đợi bằng các trung bình mẫu và  $\varepsilon_t$  bởi  $e_t$ , tạo ra ma trận  $S^*$ . Bởi vì với một mẫu chỉ có một trung bình mẫu  $(1/T) \sum_{t=1}^T x_t \varepsilon_t$  để xem xét, chúng tôi ước lượng phương sai của nó bằng cách xem xét phương sai, trong giới hạn mẫu, của trung bình ngắn hạn,  $(1/H) \sum_{j=0}^{H-1} x_{t-j} \varepsilon_{t-j}$  và phân chia theo cỡ mẫu  $T$ . Bởi vì  $S^*$  ước lượng một ma trận hiệp phương sai tiệm cận, nghĩa là ma trận hiệp phương sai của  $\sqrt{T}$  lần trung bình mẫu, yếu tố  $(1/T)$  biến mất một lần nữa trong (4.70). Lưu ý rằng nếu không có sự tự tương quan, chúng ta sẽ ước lượng phương sai của trung bình mẫu như là phương sai của mẫu  $x_t \varepsilon_t$ , chia cho  $T$ . Cuối cùng, thực tế là ước lượng sử dụng phần dư OLS  $e_t$  và thay vì phần dư không quan sát được  $\varepsilon_t$  không có kết quả tiệm cận.

#### 4.11 Minh họa: phần bù rủi ro trong thị trường ngoại hối

Một thương nhân đặt hàng ở nước ngoài phải trả tiền vào một ngày nào đó trong tương lai có thể chọn thanh toán bằng nhiều cách khác nhau. Ví dụ: xem xét một thương nhân châu Âu vào cuối tháng hiện tại mua một lượng cà phê với giá 100000 USD và sẽ phải trả tiền vào cuối tháng sau. Chiến lược thứ nhất để thu xếp tài khoản của ông ấy là mua đô la bây giờ và giữ nó trong tài khoản tiền gửi cho đến cuối tháng tới. Điều này có một kết quả rõ ràng là nhà đầu tư nhận lãi suất Mỹ và không được nhận lãi suất Châu Âu (một tháng) trong tháng này, giả định ông ta giữ số tiền USD trong một tài khoản tiền gửi của Mỹ). Chiến lược thứ hai là mua đô la tại thị trường kỳ hạn. Có một mức giá (tỷ giá) được xác định, và phải được thanh toán bằng đô la khi nó được giao vào cuối tháng sau. Tỷ giá kỳ hạn này được thỏa thuận trong giai đoạn hiện tại và phải thanh toán khi giao hàng (một tháng kể từ bây giờ). Giả sử rằng hợp đồng kỳ hạn không rủi ro (bỏ qua rủi ro vỡ nợ, thường rất nhỏ), thương nhân sẽ không phân biệt giữa hai chiến lược. Cả hai khả năng đều không có rủi ro và do đó chúng được kỳ vọng sẽ tạo ra cùng suất sinh lợi vào cuối tháng sau. Nếu không, khả năng kiếm lời do chênh lệch giá sẽ tạo ra lợi nhuận không rủi ro. Điều này ngụ ý rằng cân bằng trong chênh lệch lãi suất (tỷ giá ở châu Âu và Mỹ) và chênh lệch giữa tỷ giá kỳ hạn và tỷ giá giao ngay được gọi là điều kiện ngang bằng lãi suất có bảo hiểm (CIP).

Một khả năng thứ ba đối với thương nhân để thanh toán hóa đơn bằng đô la đơn giản là đợi đến cuối tháng sau và sau đó mua đô la Mỹ với một tỷ giá chưa xác định. Nếu giả định thông thường được đưa ra là thương nhân không ưa thích rủi ro, sẽ chỉ hấp dẫn để chấp nhận thêm rủi ro tỷ giá hối đoái nếu có kỳ vọng rằng tỷ giá giao ngay trong tương lai (tính bằng đồng euro trên đô la) thấp hơn tỷ giá kỳ hạn. Nếu đây là trường hợp, chúng ta nói rằng thị trường sẵn sàng trả phần bù rủi ro. Trong trường hợp không có phần bù rủi ro (tỷ giá kỳ hạn bằng với tỷ giá giao ngay kỳ vọng), ngang bằng lãi suất có bảo hiểm (CIP) có nghĩa là ngang bằng lãi suất không có bảo hiểm (UIP), cho biết rằng sự khác biệt về tỷ lệ lãi suất giữa hai quốc gia tương đương với sự thay đổi tương đối dự kiến trong tỷ giá. Trong phần này, chúng tôi xem xét các kiểm định về sự tồn tại của phần bù rủi ro trong thị trường hối đoái kỳ hạn, dựa trên các mô hình hồi qui.

##### 4.11.1 Ghi chú

Nhà đầu tư châu Âu có thể phòng ngừa rủi ro tiền tệ bằng cách mua vào thời điểm  $t$  số đô la cần thiết để giao hàng tại thời điểm  $t+1$  với tỷ giá  $F_t$ , tỷ giá hối đoái kỳ hạn. Như vậy,  $F_t$  là tỷ



giá tại thời điểm  $t$  với lượng đô la có thể mua và bán (thông qua một hợp đồng kỳ hạn) tại thời điểm  $t+1$ . Lãi suất không rủi ro cho châu Âu và Hoa Kỳ lần lượt là  $R_{f,t+1}^E$  và  $R_{f,t+1}^{US}$ . Đối với nhà đầu tư châu Âu, việc đầu tư vào các khoản tiền gửi của Mỹ có thể không có rủi ro thông qua việc nghiệp vụ tự bảo hiểm (hedging) trên thị trường hối đoái kỳ hạn. Tức là, một khoản đầu tư không rủi ro cho nhà đầu tư người Đức sẽ đem lại

$$\bar{R}_{f,t+1}^{US} + \log \bar{F}_t - \log \bar{S}_t, \quad (4.71)$$

trong đó  $S_t$  là tỷ giá (hối đoái) giao ngay hiện tại. Để tránh các cơ hội kinh doanh chênh lệch giá không rủi ro (và lợi nhuận không giới hạn cho các nhà đầu tư), lợi nhuận này sẽ bằng lợi nhuận không rủi ro đối với tiền gửi của châu Âu, nghĩa là nó phải bằng

$$R_{f,t+1}^E - R_{f,t+1}^{US} = \log F_t - \log S_t. \quad (4.72)$$

Phía bên tay phải của (4.72) được gọi là (phần âm của) khoản khấu trừ (hay khoản chiết khấu), trong khi phía bên tay trái được gọi là chênh lệch lãi suất. Điều kiện (4.72) được biết đến như là điều kiện ngang bằng lãi suất có bảo hiểm và là điều kiện không có kinh doanh chênh lệch giá đơn thuần, do đó gần như chắc chắn thỏa mãn trong thực tế (nếu chi phí giao dịch là không đáng kể).

Một khoản đầu tư thay thế tương ứng với một khoản đầu tư vào tiền gửi của Hoa Kỳ mà không có phòng ngừa nguy cơ rủi ro tiền tệ. Lợi tức của đầu tư rủi ro này là

$$R_{f,t+1}^{US} + \log S_{t+1} - \log S_t, \quad (4.73)$$

giá trị kỳ vọng bằng (4.71) nếu

$$E_t\{\log S_{t+1}\} = \log F_t \quad \text{or} \quad E_t\{s_{t+1}\} = f_t,$$

trong đó các chữ cái nhỏ biểu thị log của các chữ viết hoa, và  $E_t\{.\}$  biểu thị kỳ vọng có điều kiện cho tất cả các thông tin có sẵn tại thời điểm  $t$ . Đồng thức  $E_t\{s_{t+1}\} = f_t$  cùng với ngang bằng lãi suất có bảo hiểm ám chỉ điều kiện ngang bằng lãi suất không có bảo hiểm, cho thấy sự chênh lệch lãi suất giữa hai quốc gia bằng sự thay đổi tỷ giá hối đoái kỳ vọng, tức là

$$R_{f,t+1}^E - R_{f,t+1}^{US} = E_t\{\log S_{t+1}\} - \log S_t. \quad (4.74)$$

Nhiều mô hình kinh tế vĩ mô sử dụng điều kiện UIP này. Một trong những hậu quả của nó là một quốc gia nhỏ không thể kiểm soát mức lãi suất nội địa và tỷ giá của nó. Phần tiếp theo sẽ nhằm trả lời câu hỏi liệu có giữ được ngang bằng lãi suất không có bảo hiểm, tức là liệu phân bù rủi ro trên thị trường hối đoái kỳ hạn có tồn tại hay không.

Lý do tại sao tỷ giá giao ngay tương lai kỳ vọng  $E_t\{s_{t+1}\}$  có thể khác so với tỷ giá kỳ hạn  $f_t$  là vì có sự tồn tại của phân bù rủi ro. Có thể thị trường sẵn sàng trả phân bù rủi ro để đổi lại rủi ro tỷ giá trong (4.73). Trong trường hợp không có phân bù rủi ro, việc phòng ngừa rủi ro tiền tệ là miễn phí và bất kỳ nhà đầu tư nào cũng có thể hoàn toàn loại bỏ rủi ro tỷ giá của mình mà không phải trả phí. Bởi vì phân bù rủi ro dương cho một nhà đầu tư châu Âu có nghĩa là một nhà đầu tư Hoa Kỳ có thể phòng ngừa rủi ro tỷ giá hối đoái đối với đồng euro khi nhận

được chiết khấu, sẽ là thông thường khi giả định rằng không có nhà đầu tư nào phải trả phần bù rủi ro. Trong trường hợp này, thị trường ngoại hối thường được gọi là (không có rủi ro) 'hiệu quả' (xem Taylor, 1995).

Lưu ý rằng phần bù rủi ro được xác định là sự khác biệt giữa log của tỷ giá giao ngay tương lai kỳ vọng và log của tỷ giá kỳ hạn. Việc thêm lôgarít vào có một mục tiêu quan trọng nhằm thể hiện tỷ giá hối đoái trong một hoặc một loại tiền tệ khác không còn là vấn đề nữa. Trong trường hợp lôgarít, điều này không còn là vấn đề bởi vì

$$E_t\{\log S_{t+1}^{-1}\} - \log F_t^{-1} = -E_t\{\log S_{t+1}\} + \log F_t$$

#### 4.11.2 Kiểm định phần bù rủi ro của thị trường trong một tháng

Một cách tiếp cận để kiểm tra sự hiện diện của phần bù rủi ro là dựa trên một khuôn khổ hồi qui đơn giản. Trong phần này, chúng ta sẽ thảo luận các bài kiểm định về sự có mặt của phần bù rủi ro trong thị trường kỳ hạn 1 tháng sử dụng dữ liệu hàng tháng. Nghĩa là, tần suất lấy mẫu tương ứng với độ dài của hợp đồng. Các kết quả thực nghiệm sẽ được trình bày cho các kỳ hạn 1 tháng đối với tỷ giá hối đoái US\$/€ và US\$/£ Sterling, sử dụng dữ liệu hàng tháng từ tháng 1 năm 1979 đến tháng 12 năm 2001. Tỷ giá hối đoái tiền-euro (pre-euro) được dựa trên Mác (Mark) Đức. Việc sử dụng dữ liệu hàng tháng để kiểm tra phần bù rủi ro trên thị trường kỳ hạn 3 tháng sẽ được thảo luận trong phần tiếp theo.

Giả thuyết rằng không có phần bù rủi ro có thể được viết như sau

$$H_0: E_{t-1}\{s_t\} = f_{t-1}. \quad (4.75)$$

Một cách đơn giản để kiểm tra giả thuyết này là khai thác kết quả đã biết rằng sự khác biệt giữa một biến ngẫu nhiên và kỳ vọng có điều kiện của nó với một tập hợp thông tin nhất định là không tương quan với bất kỳ biến nào từ tập hợp thông tin này, tức là

$$E\{(s_t - E_{t-1}\{s_t\})x_{t-1}\} = 0 \quad (4.76)$$

với bất kỳ  $x_{t-1}$  nào được biết đến vào thời điểm  $t-1$ . Từ đó chúng ta có thể viết mô hình hồi qui như sau

$$s_t - f_{t-1} = x'_{t-1}\beta + \varepsilon_t, \quad (4.77)$$

trong đó  $\varepsilon_t = s_t - E_{t-1}\{s_t\}$ . Nếu  $H_0$  là chính xác và nếu  $x_{t-1}$  được biết ở thời gian  $t-1$  thì  $\beta = 0$ . Do đó,  $H_0$  dễ dàng được kiểm định bằng cách kiểm tra xem liệu  $\beta = 0$  cho một số các biến  $x_{t-1}$ . Phần dưới đây chúng ta sẽ chọn các phần tử trong  $x_{t-1}$  một hằng số và khoản khấu trừ  $s_{t-1} - f_{t-1}$ .

Bởi vì  $s_{t-1} - f_{t-2}$  được quan sát trong giai đoạn  $t-1$ ,  $\varepsilon_{t-1}$  cũng là một phần của tập hợp thông tin tại thời điểm  $t-1$ . Do đó, (4.76) cũng ngụ ý rằng dưới giả thiết  $H_0$ , các sai số trong (4.77) không có tự tương quan. Tự tương quan trong  $\varepsilon_t$  là một dấu hiệu cho sự tồn tại của phần bù rủi ro. Lưu ý rằng giả thuyết không ngụ ý bất cứ điều gì về phương sai của  $\varepsilon_t$ , điều này gợi ý rằng áp dụng phương sai không thay đổi có thể không thích hợp và sai số chuẩn phương sai thay đổi nhất quán có thể được áp dụng.

Dữ liệu sử dụng<sup>6</sup> được lấy từ Datastream và bao gồm khoảng thời gian từ tháng 1 năm 1979 đến tháng 12 năm 2001. Chúng tôi sử dụng tỷ giá US\$/€ và tỷ giá US\$/£ được minh họa trong Hình 4.5. Từ hình này, chúng ta có thể suy ra điểm yếu của đồng USD vào đầu những năm 1980 và 2000/2001. Trong Hình 4.6, khoản khấu trừ hàng tháng  $s_t - f_t$  được tính toán cho cả hai tỷ giá hối đoái. Thông thường, giá trị tuyệt đối của khoản khấu trừ nhỏ hơn 1%. Đối với đồng euro, tỷ giá giao ngay đô la tại hầu như tất cả các tháng dưới tỷ giá kỳ hạn, có nghĩa là, theo ngang bằng lãi suất có bảo hiểm, lãi suất danh nghĩa của Mỹ vượt quá lãi suất danh nghĩa của châu Âu. Chỉ trong giai đoạn 1993-1994 và cuối năm 2001 diễn ra điều ngược lại.

Tiếp theo, phương trình (4.77) được ước lượng bởi OLS lấy  $x_{t-1} = (1, s_{t-1} - f_{t-1})'$ . Các kết quả cho tỷ giá US\$/£ được trình bày trong Bảng 4.12. Bởi vì khoản khấu trừ có các thuộc tính của một biến phụ thuộc trễ ( $s_{t-1} - f_{t-1}$  tương quan với  $\varepsilon_{t-1}$ ), kiểm định Durbin-Watson là không thích hợp. Cách đơn giản nhất là sử dụng kiểm định Breusch-Godfrey, dựa trên hàm hồi qui bổ sung của  $e_t$  and  $e_{t-1}$ ,  $s_{t-1} - f_{t-1}$  và một hằng số (xem phần trên) và sau đó lấy<sup>7</sup>  $TR^2$ . Chúng ta có thể kiểm định tự tương quan bậc cao hơn bằng cách thêm vào các độ trễ khác như  $e_{t-2}$  và  $e_{t-3}$ . Bằng cách này, giả thuyết  $H_0$  rằng không có tự tương quan tự trễ được kiểm định đối với các giả định thay thế của tự tương quan bậc nhất và (cho đến) bậc 12, với thống kê kiểm định là 0,22 và 10,26. Với 5% giá trị tới hạn là 3,84 và 21,0 (tương ứng với  $\chi^2_1$  và  $\chi^2_{12}$ ), điều này không ám chỉ việc bác bỏ giả thiết  $H_0$ . Kiểm định t (t-statistics) trong hồi qui cho thấy rằng tính chẵn khác biệt đáng kể so với không, trong khi khoản khấu trừ có hệ số dương đáng kể. Một kiểm định đồng thời trên hai giới hạn  $\beta = 0$  dẫn đến một thống kê f bằng 7,74 ( $p = 0,0005$ ), để giả thuyết  $H_0$  rằng không có phần bù rủi ro bị bác bỏ. Các con số này hàm ý rằng nếu lãi suất danh nghĩa của Anh vượt quá mức lãi suất của Mỹ, theo đó khoản khấu trừ  $s_{t-1} - f_{t-1}$  vượt quá 0,16% (ví dụ vào đầu thập niên 90), người ta thấy rằng  $E_{t-1} \{s_t\} - f_{t-1}$  dương. Như vậy, các nhà đầu tư Anh có thể bán bảng Anh trên thị trường kỳ hạn ở mức 1,75 USD, trong khi tỷ giá giao ngay dự kiến là 1,77 USD. Các nhà nhập khẩu Anh muốn phòng ngừa rủi ro tỷ giá cho các đơn đặt hàng ở Mỹ phải trả phí phần bù rủi ro. Mặt khác, các nhà kinh doanh Hoa Kỳ có lợi từ điều này; họ có thể phòng ngừa rủi ro tiền tệ và lĩnh tiền mặt (!) phần bù rủi ro cùng một lúc.<sup>8</sup>

<sup>6</sup> Dữ liệu cho minh họa này có sẵn trong FORWARD2.

<sup>7</sup> Dưới đây chúng tôi sử dụng số lượng các quan sát hiệu quả trong các hàm hồi qui phụ để xác định T trong  $TR^2$ .

<sup>8</sup> Không có vấn đề gì chủ yếu với phần bù rủi ro bị âm. Mặc dù điều này có nghĩa là lợi nhuận kỳ vọng thấp hơn khoản đầu tư không có rủi ro nhưng lợi nhuận thực tế vẫn có thể vượt quá tỷ lệ không có rủi ro trong những tình huống đặc biệt thú vị đối với nhà đầu tư. Ví dụ, bảo hiểm tài sản trong nhà của bạn thường có lợi nhuận mong đợi âm, nhưng lại có lợi nhuận dương lớn trong trường hợp cụ thể ngôi nhà của bạn bị cháy.



Figure 4.5 US\$/EUR and US\$/GBP exchange rates, January 1979–December 2001



Figure 4.6 Forward discount, US\$/EUR and US\$/GBP, January 1979–December 2001

Table 4.12 OLS results US\$/£Sterling

| Dependent variable: $s_t - f_{t-1}$                                       |          |                |                 |
|---------------------------------------------------------------------------|----------|----------------|-----------------|
| Variable                                                                  | Estimate | Standard error | <i>t</i> -ratio |
| constant                                                                  | -0.0051  | 0.0024         | -2.162          |
| $s_{t-1} - f_{t-1}$                                                       | 3.2122   | 0.8175         | 3.929           |
| $s = 0.0315 \quad R^2 = 0.0535 \quad \bar{R}^2 = 0.0501 \quad F = 15.440$ |          |                |                 |

Kiểm định *t* được sử dụng ở trên chỉ có giá trị một cách tiệm cận khi  $\varepsilon_t$  không có sự tự tương quan, điều này được bảo đảm bởi (4.76), và nếu  $\varepsilon_t$  là phương sai không thay đổi. Kiểm định thống kê Breusch-Pagan về phương sai thay đổi có thể được tính bởi  $TR^2$  trong hàm hồi qui bổ sung của  $\varepsilon_t^2$  theo hằng số và  $s_{t-1} - f_{t-1}$ , đem lại giá trị 7,26, có nghĩa là rõ ràng bác bỏ giả thiết  $H_0$ . Việc sử dụng các sai số chuẩn phương sai thay đổi nhất quán không dẫn đến các kết luận khác nhau một cách định tính.

Theo cách tương tự, chúng ta có thể kiểm định phần bù rủi ro với tỷ giá kỳ hạn US\$/€. Kết quả của hồi qui này như sau

$$s_t - f_{t-1} = -0.0023 + 0.485 (s_{t-1} - f_{t-1}) + e_t, \quad R^2 = 0.0015$$

$$(0.0031) \quad (0.766)$$

$$BG(1) = 0.12, \quad BG(12) = 14.12.$$

Ở đây, BG(h) biểu thị thống kê kiểm định Breusch-Godfrey cho sự tự tương quan đến bậc h. Đối với tỷ giá USD/€ không tìm thấy phần bù rủi ro: các hệ số hồi qui khác không không có ý nghĩa và giả thuyết về việc không có tự tương quan không bị bác bỏ.

#### 4.11.3 Kiểm định phần bù rủi ro sử dụng mẫu lặp (overlapping sample)

Các khoản mục trước đó bị giới hạn trong phân tích của thị trường hối đoái kỳ hạn một tháng. Tất nhiên, các thị trường kỳ hạn tồn tại với các kỳ đáo hạn khác, ví dụ 3 tháng hoặc 6 tháng. Trong phần này chúng ta sẽ chú ý đến câu hỏi làm cách nào để các kỹ thuật được thảo luận trong phần trước có thể được sử dụng để kiểm tra sự có mặt của phần bù rủi ro trong thị trường kỳ hạn 3 tháng. Tần suất quan sát vẫn là một tháng.

Giả sử log của giá cả một hợp đồng kỳ hạn 3 tháng là  $f_t^3$ . Giả thuyết  $H_0$  rằng không có phần bù rủi ro có thể được thể hiện trong công thức sau

$$H_0: E_{t-3}\{s_t\} = f_{t-3}^3. \quad (4.78)$$

Sử dụng những lập luận tương tự như trước, một mô hình hồi qui tương tự như (4.77) có thể được viết lại như sau

$$s_t - f_{t-3}^3 = x'_{t-3}\beta + \varepsilon_t, \quad (4.79)$$

trong đó  $\varepsilon_t = s_t - E_{t-3}\{s_t\}$ . Nếu  $x_{t-3}$  được quan sát tại thời điểm  $t-3$ , vector  $\beta$  trong (4.79) phải bằng không theo giả thiết  $H_0$ . Sử dụng OLS để ước lượng các thông số trong (4.79),  $x_{t-3} = (1, s_{t-3} - f_{t-3}^3)'$  với các kết quả sau cho tỷ giá US\$/£:

$$s_t - f_{t-3}^3 = -0.014 + 3.135 (s_{t-3} - f_{t-3}^3) + e_t, \quad R^2 = 0.1146$$

(0.004) (0.529)

$$BG(1) = 119.69, \quad BG(12) = 173.67.$$

và cho tỷ giá US\$/€

$$s_t - f_{t-3}^3 = -0.011 + 0.006 (s_{t-3} - f_{t-3}^3) + e_t, \quad R^2 = 0.0000$$

(0.006) (0.535)

$$BG(1) = 130.16, \quad BG(12) = 177.76.$$

Những kết quả này dường như đề xuất sự có mặt một cách rõ ràng của phần bù rủi ro ở cả hai thị trường: kiểm định Breusch-Godfrey chỉ ra sự tồn tại mạnh mẽ của tự tương quan, trong khi các hệ số hồi qui cho thị trường ngoại hối USD/£ đều có ý nghĩa cao. Tuy nhiên, những kết luận này không chính xác.

Giả định rằng các sai số thể hiện việc không có tự tương quan dựa trên quan sát rằng (4.76) giữ  $x_{t-1} = \varepsilon_{t-1}$  sao cho  $\varepsilon_{t+1}$  và  $\varepsilon_t$  không tương quan. Tuy nhiên, kết quả này chỉ có giá trị nếu tần suất của dữ liệu trùng với thời hạn của hợp đồng. Trong trường hợp hiện tại, chúng tôi có dữ liệu hàng tháng cho các hợp đồng 3 tháng. Công thức tương tự của (4.76) bây giờ là

$$E\{(s_t - E_{t-3}\{s_t\})x_{t-3}\} = 0 \text{ for any } x_{t-3} \text{ known at time } t-3. \quad (4.80)$$

Do đó, điều này hàm ý rằng  $\varepsilon_t$  và  $\varepsilon_{t-j}$  ( $j = 3, 4, 5, \dots$ ) không tương quan nhưng không ngụ ý rằng  $\varepsilon_t$  và  $\varepsilon_{t-1}$  hoặc  $\varepsilon_{t-2}$  không tương quan. Ngược lại, các sai số này có khả năng tương quan cao.

Xem xét một trường hợp minh họa khi (log) tỷ giá hối đoái được tạo ra bởi các bước đi ngẫu nhiên (random walk)<sup>9</sup>, tức là  $s_t = s_{t-1} + \eta_t$  trong đó  $\eta_t$  độc lập và có cùng giá trị trung bình bằng không và phương sai  $\sigma_\eta^2$  và không có phần bù rủi ro,  $f_{t-3}^3 = E_{t-3}\{s_t\}$ . Sau đó, có thể dễ dàng chỉ ra rằng

$$\varepsilon_t = s_t - E_{t-3}\{s_t\} = \eta_t + \eta_{t-1} + \eta_{t-2}.$$

Do đó, sai số  $\varepsilon_t$  được mô tả bởi mô hình tự tương quan trung bình trượt hoặc theo bậc 2. Khi (log) tỷ giá hối đoái không phải là các bước đi ngẫu nhiên, sai số  $\varepsilon_t$  sẽ bao gồm 'tin tức' từ các khoảng thời gian  $t$ ,  $t-1$  và  $t-2$ , và do đó  $\varepsilon_t$  là trung bình trượt ngay cả trong trường hợp tổng quát hơn. Sự tự tương quan này là do vấn đề mẫu lặp, trong đó tần suất quan sát (hàng tháng) cao hơn tần suất dữ liệu (hàng quý). Nếu chúng ta kiểm tra xem sự tự tương quan có vượt quá hai độ trễ hay không, nghĩa là liệu  $\varepsilon_t$  có tương quan với  $\varepsilon_{t-3}$  đến  $\varepsilon_{t-12}$  hay không, chúng ta có thể chạy hồi qui phần dư của OLS  $\varepsilon_t$  theo  $\varepsilon_{t-3}$ , ...,  $\varepsilon_{t-12}$  và các biến hồi qui từ (4.79). Kết quả này dẫn đến kết quả thống kê kiểm định Breusch-Godfrey là 7,85 và 9,04, cả hai đều không có ý nghĩa đối với phân phối Chi-squared với 10 bậc tự do.

Thực tế là hai sự tự tương quan đầu tiên của các sai số trong các hồi qui ở trên khác không hàm ý rằng các kết quả hồi qui không cung cấp nhiều thông tin về sự tồn tại của phần bù rủi ro: sai số chuẩn được tính theo cách không chính xác, và kiểm định Breusch-Godfrey cho tự tương quan có thể bị bác bỏ bởi vì hai sự tự tương quan đầu tiên khác không, điều này không liên quan đến việc không có phần bù rủi ro. Lưu ý rằng ước lượng OLS vẫn nhất quán, ngay cả khi tồn tại sai số dịch chuyển trung bình.

Một cách để "giải quyết" vấn đề tự tương quan đơn giản là bỏ bớt 2/3 thông tin bằng cách chỉ sử dụng các quan sát với các khoảng ba tháng. Điều này là không thỏa mãn, vì phải bỏ mất thông tin và do đó có khả năng làm giảm năng lực của kiểm định. Có thể nghĩ đến hai lựa chọn khác: (i) sử dụng GLS để (hy vọng là) ước lượng mô hình một cách hiệu quả hơn, và (ii) sử dụng OLS trong khi tính toán các sai số chuẩn điều chỉnh (Newey-West). Thật không may, lựa chọn đầu tiên không thích hợp trong trường hợp này bởi vì dữ liệu được chuyển đổi sẽ không đáp ứng các điều kiện cho tính nhất quán và GLS sẽ không nhất quán. Điều này là do thực tế rằng hồi qui  $s_{t-3} - f_{t-3}^3$  tương quan với các sai số trễ.

Do đó chúng tôi sẽ xem xét các kết quả ước lượng OLS một lần nữa, bằng cách tính các sai số chuẩn HAC. Lưu ý rằng  $H = 3$  là đủ. Nhớ lại rằng các sai số chuẩn này cho phép phương sai thay đổi. Các kết quả có thể được tóm tắt như sau: Đối với tỷ giá US\$/£ chúng ta có

$$s_t - f_{t-3}^3 = -0.014 + 3.135 (s_{t-3} - f_{t-3}^3) + e_t, \quad R^2 = 0.1146$$

[0.005] [0.663]

và đối với tỷ giá US\$/€

$$s_t - f_{t-3}^3 = -0.011 + 0.006 (s_{t-3} - f_{t-3}^3) + e_t, \quad R^2 = 0.0000$$

[0.008] [0.523]

<sup>9</sup> Chi tiết về các bước đi ngẫu nhiên được trình bày trong chương 8.

trong đó các sai số chuẩn trong ngoặc vuông là các sai số chuẩn Newey-West với  $H = 3$ . Một cách định tính thì, kết luận không thay đổi: đối với thị trường tỷ giá US\$/£ 3 tháng, ngang bằng lãi suất không có bảo hiểm bị bác bỏ. Bởi vì ngang bằng lãi suất có bảo hiểm hàm ý rằng

$$s_t - f_t = R_{f,t+1}^* - R_{f,t+1},$$

với \* biểu thị nước ngoài và tỷ giá hối đoái được đo bằng số lượng đơn vị tiền tệ trong nước đổi một đơn vị ngoại tệ, kết quả cho thấy khi mà lãi suất của Mỹ cao hơn một cách tương đối so với Anh, các nhà đầu tư Anh phải trả phần bù rủi ro cho thương nhân Mỹ. Đối với thị trường Châu Âu/Hoa Kỳ, sự tồn tại của phần bù rủi ro không được tìm thấy trong dữ liệu.