

DỰ BÁO HOẠT ĐỘNG THƯƠNG MẠI NỘI ĐỊA DỰA TRÊN MÔ HÌNH TỰ HỒI QUY VEC TƠ (VAR)

Mã số: 98.1GEMg.12

Cao Tuấn Khanh
Trưởng Đại học Thương mại
Email: caotuan Khanh2000@yahoo.com

Ngày nhận: 30/08/2016 Ngày nhận lại: 05/09/2016 Ngày duyệt đăng: 19/09/2016

Nghiên cứu này tập trung dự báo hoạt động thương mại nội địa trong phạm vi lãnh thổ Việt Nam (đại diện là doanh số bán lẻ hàng hóa và dịch vụ tiêu dùng) dựa trên mô hình Tự hồi quy Vec tơ (VAR), trong đó có tính đến các yếu tố chính tác động chính như: (i) thu nhập bình quân đầu người; (ii) độ mở của nền kinh tế; (iii) lạm phát và (iv) cung ứng tiền.

Từ khóa: VAR, thương mại nội địa.

1. Giới thiệu

Dự báo sự phát triển của hoạt động thương mại có vai trò quan trọng trong việc hoạch định và phối hợp các chính sách vĩ mô liên quan đến thương mại, đầu tư và hướng phát triển thị trường hàng hóa và dịch vụ nội địa trong tương lai. Trên thế giới, có khá nhiều phương pháp và mô hình định lượng được sử dụng để dự báo các chỉ tiêu vĩ mô như vậy bao gồm: mô hình trung bình trượt (ARIMA), mô hình tự hồi quy Vec tơ (VAR) và các biến thể, mô hình kinh tế lượng vĩ mô (Macro Econometrics Model) hay mô hình mạng thần kinh nhân tạo.

Trong số đó, mô hình VAR và các dạng biến thể của nó thường được các nhà nghiên cứu sử dụng khi phân tích cũng như dự báo trong trung hạn. Với một số ít biến trong mô hình nhưng tính tin cậy của dự báo thu được lại tốt hơn so với một mô hình hệ phương trình đồng thời công kênh nhiều biến và nhiều phương trình. Mặt khác, mô hình này cũng có ưu điểm trong việc xây dựng các kịch bản chính sách.

2. Cơ sở lý thuyết mô hình VAR

Trước năm 1980, nhiều NHTW trên thế giới và các cơ quan dự báo kinh tế vĩ mô chủ yếu sử dụng mô hình hệ phương trình đồng thời trong các nghiên cứu về chu kỳ kinh tế và cung tiền. Tuy nhiên, nhược điểm của ước lượng hệ phương trình đồng thời bằng phương pháp ước lượng bình phương nhỏ nhất là kết quả ước lượng bị chệch mà

không vững (nghĩa là ngay khi kích thước mẫu tiến ra vô cùng thì giá trị ước lượng cũng không hội tụ về giá trị thực).

Với sự ra đời của lý thuyết tăng trưởng kinh tế nội sinh trong những năm 1980 thường được gọi là lý thuyết tăng trưởng kinh tế mới, các nhà kinh tế đã xem xét các yếu tố nội sinh giữa các biến kinh tế vĩ mô và nhận định hầu hết các biến kinh tế vĩ mô đều là các biến nội sinh và có mối quan hệ qua lại với nhau. Trong đó nổi bật là nghiên cứu của Sims khi tập trung vào việc phân biệt giữa những thay đổi không lường trước của các biến số như giá dầu hay lãi suất và những thay đổi đã lường trước để xác định tác động của chúng đối với các biến số vĩ mô quan trọng. Ông cho rằng hầu hết các biến số kinh tế, nhất là biến kinh tế vĩ mô đều mang tính nội sinh, nghĩa là đều có tác động qua lại với nhau. Từ đó, Sims đã phát triển mô hình Tự hồi quy Vec tơ (VAR) với nhiều biến số, trong đó các biến số đều là nội sinh và đóng vai trò như nhau để phân tích xem những cú sốc từ một biến số sẽ tác động ra sao tới các biến số vĩ mô khác. Phương pháp của ông có tác động rất lớn đối với nghiên cứu định lượng và là nền tảng cho việc ban hành chính sách kinh tế đến tận ngày nay.

Mô hình VAR là một mô hình kinh tế lượng dùng để xem xét động thái và sự phụ thuộc lẫn nhau giữa một biến số theo thời gian. Trong mô hình VAR, mỗi biến số được giải thích bằng một phương trình chứa các giá trị trễ của chính biến số

và các giá trị trễ của các biến số khác.

Mô hình VAR cấp p tổng quát có dạng:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + s_t + u_t$$

Trong đó:

- A_i là ma trận vuông cấp $m \times m$,

$i=1,2,\dots,p$; $s_t = (s_{1t}, s_{2t}, \dots, s_{mt})$

- Y bao gồm m biến ngẫu nhiên dừng;

- u_t là véc tơ các nhiễu trắng;

- s_t véc tơ các yếu tố xác định, có thể bao gồm hằng số, xu thế tuyến tính hoặc đa thức.

Viết dưới dạng toán tử trễ, ta có:

$$Y_t = (A_1 L + A_2 L^2 + \dots + A_p L^p) Y_t + s_t + u_t$$

Mô hình SVAR thông thường có dạng:

$$A y_t = C(L) y_{t-1} + B u_t$$

Trong đó:

- A là một ma trận ($G \times G$) các

hệ số đồng thời xác định các mối quan hệ đồng thời giữa G biến

trong cùng một khoảng thời gian

- $C(L) = C_0 - C_1 L - C_2 L^2 - \dots - C_p L^p$ mô tả các biến động tương tác giữa các biến (ví dụ như độ trễ).

- $u_t = [u_{1t} \dots u_{Gt}]'$ là một véc tơ nhiễu trắng ($G \times 1$) được định nghĩa bởi $E(u_t) = 0$; $E(u_t u_s') = \sum_u$ khi $t = s$ và $E(u_t u_s') = 0$ trong trường hợp khác.

- B là một ma trận ($G \times G$) xác định các mối quan hệ đồng thời trong những cú sốc cấu trúc.

- $C(1) = C_0 - C_1 - C_2 - \dots - C_p$ xác định phản ứng dài hạn của hệ thống với một cú sốc cấu trúc.

3. Ứng dụng mô hình VAR vào dự báo thương mại nội địa của Việt Nam

Cấu trúc mô hình VAR

Nghiên cứu xác định cấu trúc biến số của mô hình VAR gồm các nhân tố chính sau:

(i) Tỷ lệ thu nhập bình quân đầu người: người dân càng có thu nhập cao hơn càng có điều kiện để tiêu thụ hàng hóa và dịch vụ nhiều hơn và với giá trị lớn hơn.

(ii) Độ mở của nền kinh tế: được đo bằng tổng kim ngạch hàng hóa xuất nhập khẩu so với GDP. Tỷ lệ này càng cao cho thấy mức độ giao thương trong nước với quốc tế càng lớn.

(iii) Chỉ số giá tiêu dùng: thể hiện mức giá bình quân của hàng hóa tiêu dùng, do đó mức giá càng cao thì giá trị hàng hóa và dịch vụ tiêu thụ trong nước (theo giá danh nghĩa) cũng càng lớn.

(iv) Tổng phương tiện thanh toán: quy mô thanh toán của nền kinh tế càng lớn đồng nghĩa với giá trị hàng hóa và dịch vụ thanh toán nội địa càng cao.

Bảng 1 : Các biến số trong mô hình VARs

STT	Tên biến	Nội dung biến	Dạng biến	Độ dài mẫu
Biến nội sinh				
1	RETAIL	Doanh số bán lẻ hàng hóa và dịch vụ tiêu dùng	Logarit cơ số tự nhiên	Q4/2006 - Q2/2016
2	GDP_CAPITA	Gdp bình quân đầu người (đại diện cho thu nhập bình quân đầu người)	Logarit cơ số tự nhiên	Q4/2006 - Q2/2016
3	OPEN	Độ mở của nền kinh tế	Logarit cơ số tự nhiên	Q4/2006 - Q2/2016
4	CPI	Chỉ số giá tiêu dùng	Logarit cơ số tự nhiên	Q4/2006 - Q2/2016
5	M2	Tổng phương tiện thanh toán	Logarit cơ số tự nhiên	Q4/2006 - Q2/2016
Biến ngoại sinh				
6	c	Hệ số chặn		

Về dữ liệu sử dụng trong mô hình: dữ liệu về GDP, CPI, dân số được thu thập từ nguồn Tổng Cục thống kê; dữ liệu về xuất nhập khẩu hàng hóa được thu thập từ Tổng Cục hải quan; dữ liệu M2 được thu thập từ Ngân hàng Nhà nước Việt Nam. Tất cả các dữ liệu sử dụng để ước lượng mô hình này đều có độ dài từ quý 4/2006 đến hết quý 2/2016. Các biến số để ước lượng mô hình đều dưới dạng Logarit cơ số tự nhiên.

Số lượng trễ của mô hình: các kiểm định thống kê như LR, FPE, HQ cho kết quả kiểm tra số trễ phù hợp là 2, trong khi SC cho kết quả 1 trễ và AIC cho kết quả 2 trễ. Do đó, tác giả lựa chọn số trễ tối ưu trong mô hình là 2 (dựa trên số kiểm định thống kê nhiều hơn), cụ thể như bảng sau:

Bảng 2 : Lựa chọn độ dài trễ tối ưu

Lag	LogL	LR	FPE	AIC	SC	HQ
0	200.1074	NA	1.35e-11	-10.83930	-10.61937	-10.76254
1	448.2546	413.5786	5.66e-17	-23.23636	-21.91677*	-22.77579
2	486.5865	53.23874*	2.94e-17*	-23.97703	-21.55776	-23.13264*
3	512.6383	28.94644	3.50e-17	-24.03546*	-20.51653	-22.80726

* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

Kết quả ước lượng mô hình

	LOG(RETAIL)	LOG(GDP_CAPITA)	LOG(OPEN)	LOG(CPI)	LOG(M2)
LOG(RETAIL(-1))	0.413172 (0.21514) [1.92048]	-0.094747 (0.05051) [-1.87567]	0.217439 (0.13776) [1.57835]	0.019111 (0.06527) [0.29281]	0.069903 (0.10564) [0.66173]
LOG(RETAIL(-2))	0.085504 (0.22650) [0.37750]	0.055225 (0.05318) [1.03843]	-0.273672 (0.14504) [-1.88690]	0.156163 (0.06871) [2.27269]	-0.066053 (0.11121) [-0.59392]
LOG(GDP_CAPITA(-1))	0.710396 (0.88886) [0.79922]	1.192573 (0.20870) [5.71425]	0.486437 (0.56918) [0.85463]	0.473138 (0.26965) [1.75462]	-0.357084 (0.43645) [-0.81816]
LOG(GDP_CAPITA(-2))	-1.446415 (0.74741) [-1.93522]	-0.348833 (0.17549) [-1.98777]	-0.405244 (0.47860) [-0.84673]	-0.398769 (0.22674) [-1.75869]	0.378237 (0.36699) [1.03064]
LOG(OPEN(-1))	-0.042963 (0.25516) [-0.16838]	0.070112 (0.05991) [1.17028]	1.089488 (0.16339) [6.66809]	0.078710 (0.07741) [1.01684]	0.070208 (0.12529) [0.56038]
LOG(OPEN(-2))	-0.118381 (0.25176) [-0.47021]	-0.070927 (0.05911) [-1.19985]	-0.456883 (0.16121) [-2.83400]	-0.045117 (0.07638) [-0.59072]	0.015270 (0.12362) [0.12352]
LOG(CPI(-1))	-1.039916 (0.66234) [-1.57007]	0.310099 (0.15551) [1.99403]	0.730847 (0.42412) [1.72320]	1.148094 (0.20093) [5.71384]	-0.618263 (0.32522) [-1.90107]
LOG(CPI(-2))	1.843993 (0.59457) [3.10137]	-0.107293 (0.13960) [-0.76856]	-0.895577 (0.38073) [-2.35226]	-0.569470 (0.18037) [-3.15716]	0.646755 (0.29194) [2.21533]
LOG(M2(-1))	0.183259 (0.54247) [0.33782]	0.089804 (0.12737) [0.70506]	0.049216 (0.34737) [0.14168]	0.192491 (0.16457) [1.16966]	0.648615 (0.26636) [2.43507]
LOG(M2(-2))	0.393300 (0.47478) [0.82838]	-0.035324 (0.11148) [-0.31688]	0.058549 (0.30402) [0.19258]	-0.203663 (0.14403) [-1.41400]	0.271539 (0.23312) [1.16478]
C	-3.211784 (2.55427) [-1.25742]	-0.852360 (0.59973) [-1.42124]	1.527934 (1.63560) [0.93417]	-0.626954 (0.77488) [-0.80910]	0.625606 (1.25418) [0.49881]
R-squared	0.990168	0.999507	0.874483	0.997341	0.998535
Adj. R-squared	0.986387	0.999317	0.826208	0.996318	0.997972
Sum sq. resids	0.058279	0.003213	0.023897	0.005364	0.014051
S.E. equation	0.047345	0.011116	0.030317	0.014363	0.023247
F-statistic	261.8502	5272.078	18.11438	975.2109	1772.410
Log likelihood	66.88766	120.5022	83.38048	111.0216	93.20497
Akaike AIC	-3.020954	-5.919036	-3.912458	-5.406573	-4.443512
Schwarz SC	-2.542033	-5.440114	-3.433537	-4.927652	-3.964591
Mean dependent	13.19624	1.906994	5.069312	4.372511	14.87767
S.D. dependent	0.405781	0.425507	0.072722	0.236710	0.516197
Determinant resid covariance (dof adj.)		7.77E-18			
Determinant resid covariance		1.33E-18			
Log likelihood		498.9560			
Akaike information criterion		-23.99762			
Schwarz criterion		-21.60301			

Một số kết quả kiểm định mô hình

- Mô hình thỏa mãn yêu cầu về tính ổn định:

Root	Modulus
0.972778	0.972778
0.929364 - 0.153029i	0.941879
0.929364 + 0.153029i	0.941879
0.800641 - 0.382210i	0.887193
0.800641 + 0.382210i	0.887193
0.451315 - 0.744628i	0.870721
0.451315 + 0.744628i	0.870721
0.015752 - 0.766694i	0.766856
0.015752 + 0.766694i	0.766856
-0.519415 - 0.389186i	0.649043
-0.519415 + 0.389186i	0.649043
0.047328 - 0.424150i	0.426782
0.047328 + 0.424150i	0.426782
-0.057258 - 0.295130i	0.300634
-0.057258 + 0.295130i	0.300634

No root lies outside the unit circle.
VAR satisfies the stability condition.

- Mô hình không có tương quan chuỗi

VAR Residual Serial Correlation LM Tests
Null Hypothesis: no serial correlation at lag order h

Date: 09/25/16 Time: 22:29

Sample: 2006Q4 2016Q2

Included observations: 36

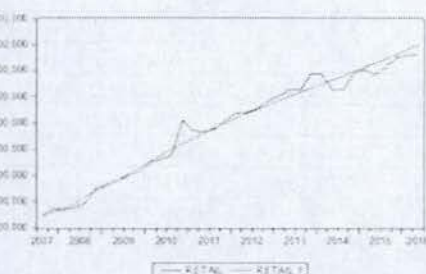
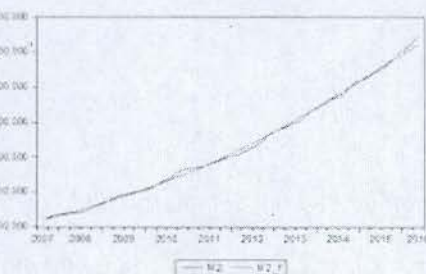
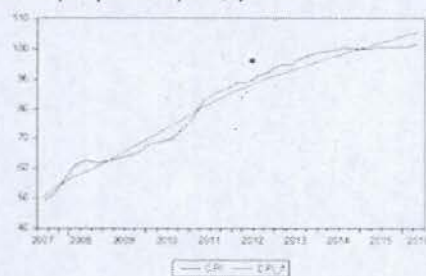
Lags	LM-Stat	Prob
1	33.36659	0.1222
2	38.50014	0.4413
3	29.99565	0.2245
4	19.74841	0.7599
5	29.43256	0.2462
6	21.62452	0.6573
7	22.14727	0.6272
8	16.25778	0.9069
9	15.49158	0.9291
10	14.66451	0.9489
11	13.12200	0.9750
12	27.30960	0.3406

Như vậy, mô hình đã vượt qua được một số kiểm định cơ bản nên có thể ứng dụng cho phân tích và dự báo.

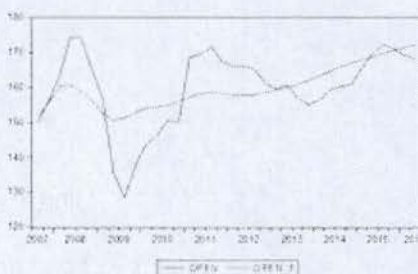
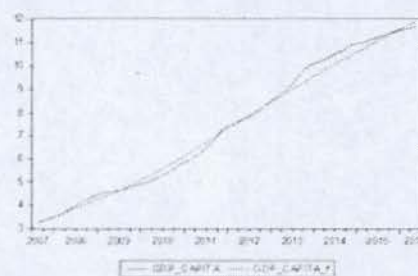
Kết quả dự báo:

- Dự báo trong mẫu

So sánh kết quả dự báo trong mẫu và diễn biến thực tế trong quá khứ tại Hình 1 cho thấy mô hình có khả năng dự báo của mô hình có thể chấp nhận được đối với biến số retail, M2, GDP_capita. Tính toán cụ thể khả năng mô phỏng của mô hình dựa trên các phương pháp MAPE, MAE, RMSE, Theil (bảng 3) cho thấy sai lệch giữa số liệu thực tế và số liệu dự báo cũng không quá lớn. Điều này chứng tỏ mô hình có thể sử dụng cho dự báo doanh số bán lẻ hàng hóa và dịch vụ tiêu dùng (retail) ngoài mẫu.



Probs from chi-square with 25 df.

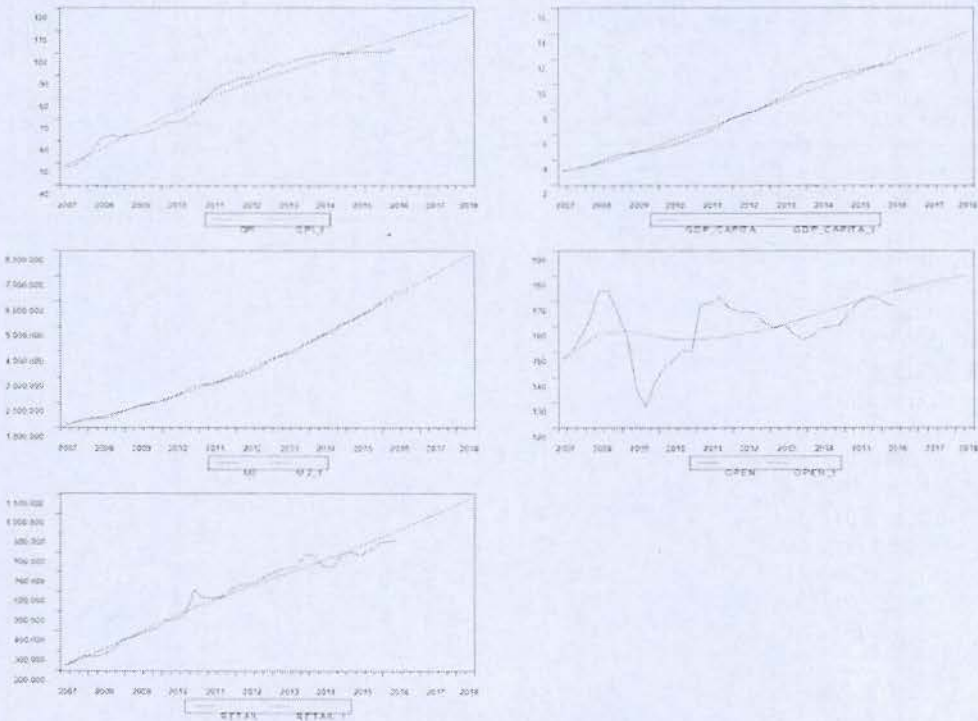


Hình 1: Kết quả dự báo trong mẫu mô hình

Bảng 3 Kết quả tính toán sai số dự báo trong mẫu của mô hình

Forecast:	CPI f	GDP_CAPITA f	M2 f	OPEN f	RETAIL f
Actual:	CPI	GDP_CAPITA	M2	OPEN	RETAIL
Forecast sample:	2007Q3 2016Q2	2007Q3 2016Q2	2007Q3 2016Q2	2007Q3 2016Q2	2007Q3 2016Q2
Root Mean Squared Error:	2.485634	0.239354	76157.32	8.625226	27357.65
Mean Absolute Error:	2.239839	0.187788	57377.03	6.888733	19635.40
Mean Abs. Percent Error:	2.801445	2.564866	1.890752	4.371348	3.261692
Theil Inequality Coefficient:	0.014818	0.015137	0.010400	0.026937	0.022061

Dự báo ngoài mẫu



Hình 2 Kết quả dự báo ngoài mẫu mô hình

Bảng 4 Kết quả tính toán sai số dự báo ngoài mẫu của mô hình

Forecast:	CPI f	GDP_CAPITA f	M2 f	OPEN f	RETAIL f
Actual:	CPI	GDP_CAPITA	M2	OPEN	RETAIL
Forecast sample:	2007Q2 2018Q2	2007Q2 2018Q2	2007Q2 2018Q2	2007Q2 2018Q2	2007Q2 2018Q2
Root Mean Squared Error:	3.195137	0.306173	75139.50	10.38487	30898.11
Mean Absolute Error:	2.890857	0.242498	56625.32	7.987971	24295.63
Mean Abs. Percent Error:	3.623736	3.337070	1.934859	5.044915	4.191307
Theil Inequality Coefficient:	0.019229	0.019610	0.010386	0.032458	0.025233

Kết quả dự báo CPI, GDP bình quân đầu người, M2 và doanh số bán lẻ hàng hóa giai đoạn 2017-2020 theo mô hình VAR

Bảng 5: Kết quả dự báo tốc độ tăng trưởng đến 2020 thực sự phù hợp và cần cải thiện để ước lượng chính xác hơn và VAR là mô hình thích hợp.

	CPI	GDP CAPITA	M2	OPEN	RETAIL
2017Q1	4.91	9.36	13.46	2.05	9.01
2017Q2	4.85	9.19	13.11	1.94	8.88
2017Q3	4.78	9.02	12.78	1.84	8.73
2017Q4	4.71	8.86	12.46	1.74	8.57
2018Q1	4.62	8.70	12.15	1.65	8.41
2018Q2	4.54	8.53	11.85	1.56	8.23
2018Q3	4.45	8.37	11.56	1.48	8.04
2018Q4	4.35	8.21	11.28	1.42	7.85
2019Q1	4.25	8.04	11.01	1.36	7.66
2019Q2	4.15	7.87	10.75	1.31	7.47
2019Q3	4.05	7.71	10.50	1.26	7.27
2019Q4	3.95	7.54	10.26	1.23	7.08
2020Q1	3.85	7.38	10.02	1.20	6.89
2020Q2	3.75	7.21	9.79	1.17	6.70
2020Q3	3.65	7.05	9.57	1.15	6.52
2020Q4	3.56	6.88	9.35	1.13	6.34

4. Mô hình OLS ước lượng doanh số bán lẻ hàng hóa và dịch vụ tiêu dùng

**Kết quả ước lượng mô hình:
Một số hạn chế của mô hình:**

Dependent Variable: LOG(RETAIL)
Method: Least Squares
Date: 09/08/16 Time: 09:58
Sample (adjusted): 2005Q4 2016Q2
Included observations: 43 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.530752	0.676378	3.741621	0.0006
LOG(GDP_POPU)	-0.314834	0.102051	-3.085051	0.0038
LOG(OPENESS)	-0.386751	0.111348	-3.473352	0.0013
LOG(CPI)	1.235495	0.206874	5.972218	0.0000
LOG(M2)	0.526387	0.091723	5.738864	0.0000
R-squared	0.986696	Mean dependent var	13.04590	
Adjusted R-squared	0.985296	S.D. dependent var	0.534458	
S.E. of regression	0.064809	Akaike info criterion	-2.525787	
Sum squared resid	0.159610	Schwarz criterion	-2.320996	
Log likelihood	59.30441	Hannan-Quinn criter.	-2.450266	
F-statistic	704.5691	Durbin-Watson stat	0.548617	
Prob(F-statistic)	0.000000			

+ Hệ số ước lượng đối với biến số log (GDP_popu) và log(openess) mang dấu âm là không phù hợp xét về mối quan hệ kinh tế.

+ Mô hình có R2 cao (0.986696) nhưng vẫn tồn tại tự tương quan (thể hiện qua thống kê Durbin-Watson), do đó các kết quả tính toán khoảng tin cậy có thể bị sai lệch.

Với các lý do trên cho thấy mô hình không

Kết luận:

Xác lập một mô hình hồi quy càng nhiều biến tưởng như là càng chính xác nhưng lại gặp nhiều khó khăn về cơ sở dữ liệu theo dãy thời gian đủ dài; tin cậy và thống nhất. Nghiên cứu cho thấy việc vận dụng pooled OLS hoặc FEM, REM để phân tích mức độ, chiều hướng tác động của các biến quan sát đến tổng mức luân chuyển hàng hóa bán lẻ và doanh thu dịch vụ bán lẻ xã hội là không thực sự phù hợp. Nghiên cứu này đã vận dụng và kiểm định mô hình VAR cho thấy tính phù hợp, khả năng thích và khả năng dự báo. ♦

1. H.Mano, R.Oliver (1993), *Assessing Dimensionality and Structure of the Consumption Experience*, J. Consumer Research Vol. 20.

2. N.Shelth, B. Newman, B.Gross (1991), *Why we buy what we buy: "Theory and Prognostic Models of Consumption Values"*, J of Business Research, Vol 22.

3. S.Amstrong (2001), *Principles for Forecasting: A Handbook for Researchers and Practitioners*, Klumer Academic Publisher.

4. L.Hughes (2002), *Econometrics Take Root*, J. of Advertising Age, Ang.5.2002.

5. G.lilien, P. Kotler, K. Moorthy (1992), *Marketing Moelels*, Upper Saddle River, Prentice Hall, New Jersey.

6. Nguyễn Bách Khoa, Cao Tuấn Khanh (2014), *Marketing Thương mại*, NXB Thống kê.

7. Nguyễn Bách Khoa (1992), *Bài tập Tổ chức và Kỹ thuật Thương nghiệp*, NXB Thống kê.

Summary

The study focuses on prognosing the retailing saile of consumption goods and serice in Vietnam domestic markets based on VAR model with the main affecting factors such as: average GDP/persion, the economy openness.