doi:10.21311/001.39.9.37

**Fitting and Prediction of Dynamic Coupling Effect Between Online and Offline Total Retail Sales of Consumer Goods**

**Xiaoling Song**  
*Business School, Beijing Language and Culture University, Beijing 100083, China*

**Abstract**  
The seasonal data of offline and online total retail sales of consumer goods (TRSCG) and online market sales between 2008 and 2015 in China were selected. Then a vector error correction model was built for fitting and prediction, and used to empirically study the dynamic coupling effect between online and offline commodity retail markets. We found online sales tended to gradually invade offline sales from the perspective of retail sales of customer goods. Nevertheless, through fitting and prediction, we found a significant dynamic coupling effect between online and offline sales. In the long run, when the online retail sales of customer goods increased by 1% and other conditions were unchanged, the retail sales of offline customer goods would rise by 0.092402%.

**Keywords:** Total retail Sales of Offline Customer Goods, Online Market Retail Sales, Vector Error Correction Model, Dynamic Coupling Effect.

1. **INTRODUCTION**

As one of the "three carriages" that drive the growth of GDP, consumption acts as an engine to promote the long-term steady economic development. After the implementation of National 12th Five-Year Plan in 2009, China planned to expand domestic demand and thereby promote the GDP growth through intensified consumption. The new data issued by National Bureau of Statistics show that the GDP in 2015 increased by 6.9% as per comparable prices. Moreover, as for the overall situation of national economy in 2015, the investments of fixed assets dropped more rapidly, imports and exports both declined year-on-year, but market sales expanded quickly. Total retail sale of consumer goods (TRSCG) is always a major consumption index and has gained wide attention. The TRSCG in 2015 increased by 10.7% year-on-year (the real increasing rate after exclusion of price factors was 10.6%).

Internet has flourished in China following the perfection of network infrastructure and the rapid popularization of mobile facilities. As the carrier of E-commerce, the Internet has developed vigorously in China within only a few years. Statistics show that the E-commerce sales in China was only 26 billion RMB in 2006, but the physical commodity online retail sales rose by 31.6% in 2015 year-on-year, accounting for 10.8% of TRSCG.

TRSCG is defined as the sum of objective non-production and non-management commodities traded from an enterprise (organization) to individuals or social groups and includes the income from the provision of catering services. The targets of TRSCG investigated by National Bureau of Statistics include the legal enterprises, industrial activity organizations and individuals engaged in commodity retail activities or provision of catering services. TRSCG reflects the developing scale and rate of the commodity market from the perspective of physical consumption. This index can be used to analyze and judge the general running situation and structural change of the commodity market and helps the government with market regulation and formulation of relevant policies and measures.

Online retail sale is defined as the sum of commodity and service retail sales obtained through public online transaction platforms (including self-built websites and third-party platforms). Here the commodities and services include physical commodities and non-physical commodities (e.g. virtual commodity, services). The statistics of TRSCG include the online retail sales of physical commodities, but exclude the online retail sales of non-physical commodities. The proportion of online retail sales in TRSCG exceeded 10% since 2014. Moreover, the rapid development of the Internet has brought online shopping to millions of families. In 2015, besides the usual promotion methods such as discount and coupons, there was a new round of updating in product quality and service quality. Also the participation of branded commodities and overseas imported commodities updated the overall style and promoted the expansion of the overall market scale. For instance, the “11•11” shopping and promotion day in 2015 witnessed a daily turnover up to 91.217 billion RMB, which implied we should never ignore the motivation of consumption by online shopping. Moreover, all E-commerce companies dribble over the vast market in rural areas and make endeavor to develop rural E-commerce. It is believed that sooner or later online retail would severely impact TRSCG.
2. LITERATURE REVIEW

Western researchers started early to study E-commerce that is strongly associated with online shopping, but focused on how to build E-commerce trade models and the effects of E-commerce on economy. For instance, Soon-Yong Choi et al. clarified the concept and localization of E-commerce from the economic perspective and elaborated its roles in economic development (Choi, Stahl, Andrew and Whinston, 1996). In the 21st century, the research on E-commerce was systematic and model-based. Some representative models include Nagi&Wat model, Aufman&Wden model, Urbacaeawski model, and Turban model, which are multilevel E-commerce development models based on technology, application, security, organization, economy and strategies.

The early studies on TRSCG in China were focused on the influence factors. With TRSCG as the dependent variable, the independent variables involved gross national income, per capita household income, net currency investment, CPI, taxation, national fiscal expenditure, permanent population, saving deposits of residents, and investment in fixed assets. The methods included least square method and factor model (Xu, Jia, Song and Chen, 2008), autoregressive moving average model (ARMA) (Wang and Wang, 2014), vector autoregression (VAR), Granger causality test, pulse response and variance decomposition (Fang, 2009).

Along with the rapid development of online shopping, the research on social commodity retails started to involve online shopping in recent years. By building a power function model, Zhao analyzed online retail sales and TRSCG and found the expansion of online retail market sale scale had a diminishing effect on TRSCG (Zhao, 2013). This diminishing effect was realized by the retail price variation, rather than the variation of non-consumption scale (Zhao, 2013). Based on the unified theory of acceptance and use of technology (UTAUT), Tan et al. included relevant variables and empirically analyzed the influence factors on online shopping (Tan, Zhang and Zeng, 2014). They found the positive influence factors on the adoption of online shopping included performance expectation, contributing factors, and consumer innovation, while the negative influence factor was the perceived risk, while endeavor expectation and social influence had no effect (Tan, Zhang and Zeng, 2014). Moreover, the three moderating variables (frequency of online shopping, monthly disposable income, consumer innovation) more or less affected the intervariable relationships (Tan, Zhang and Zeng, 2014). Theoretically, it was proved that online shopping accelerated the increment of TRSCG, provided more positions for improvement of citizen purchasing power and advocated the perfection, thereby increasing TRSCG (Wu, 2015). Ni et al. empirically analyzed the differences in the characteristics of consumer followed shopping channels under different Internet use situations, and validated the relationships between Internet utilization and the preference of online shopping channels (Ni, Li and Li, 2015).

The development of E-commerce has brought severe challenges to the traditional retailing in China. Online turnover also tended to gradually invade offline turnover from the perspective of customer goods retail sales. Thus, much research has been conducted to uncover the relationship between online and offline retail markets. The representative views are listed below. (1) The theory of impact and challenge. Offline retail is called the traditional retail and is challenged by online retail. The conflict and integration between online and offline retails are altering the competition patterns of the retailing industry in China (Zhou, 2012). (2) The theory of co-development. As reported, the price difference between online and offline shopping will persist. Online shopping expands with the enlargement of market size, while the online and offline co-development is a major trend (Du and Liu, 2014). (3) The theory of complementary advantages The advantages between the Internet E-commerce mode and the traditional trade mode are complementary and cannot be completely replaced. O2O, the novel E-commerce pattern of online-offline collaboration will be the mainstream in the future trade patterns (Zhang, 2014).

The existing research on offline and online customer goods retails provides clues and methods for this study. Different from previous studies about the influence factors on TRSCG, here we target at the dynamic coupling effect between online and offline commodity retail sales.

The coupling effect is initially a physical concept and used to reveal the relationship of close cooperative interaction between two systems or more. Through the grey relational analysis, Wu et al. built a coupling evaluating index system and established a coupling association model of relevant variables (Wu, Zhao, Guan and Wang, 2016). Together with the theory of coupling coordination, Zhu et al. (Zhu and Li, 2015) and Luo et al. (Luo, He and Mao, 2013) built coupling coordinating degree evaluating index systems for comparative analysis. The above studies well uncover the intervariable interactions and are focused on the intervariable current relationships.

The VAR model regards each endogenous variable as a function of the lagged value of all endogenous variables and well reveals the mutual dynamic influences of variables and is close to the characteristic of double-variable coupling effect. Liu used VAR to empirically analyze the coupling effects among economic growth, industrial development and labor employment in China (Liu, 2010). Zhao used VAR to analyze the coupling effect between population structure and economic development in Hebei Province, China (Zhao, 2011).
Li and Qiu used VAR to study the coupling effect between regional financial agglomeration and innovative entrepreneurship (Li, 2016). They analyzed the interactions among multiple variables, explored how one variable was affected by the lag of other variables, and especially probed into the dynamic priority between variables (Li, 2016). VAR endows the coupling effect analytic model with dynamic property. However, the traditional VAR requires every variable should be steady and otherwise spurious regression would occur. Information loss would occur if the variables were simply differentiated to guarantee their stationeriness. However, with the development of the cointegration theory, for non-stationary time series, we can directly build a vector error correction (VEC) model as long as there is a cointegration relationship between variables (Gao, 2009).

Together with previous methods about the influence factors on TRSCG, we built a VEC model for fitting and prediction, and empirically studied the dynamic coupling effect between online and offline commodity retail markets. On this basis, we provide some special suggestions on updating the social commodity market in China.

3. DATA SELECTION, PROCESSING AND STATIONARITY TEST

3.1. Data Selection and Processing

We collected the online monitoring data between the first season of 2008 and the second season of 2015 from iResearch, involving 30 seasons of online shopping retail sales. We used these data as the proxy variables revealing the developing degree of online social commodity retail market, marked as ONS. The turnover of this online shopping market included the summarization of online B2C (trade retail modes where enterprises sell products and services to consumers face-to-face, including Tmall and JD.com) and C2C (consumer-consumer E-commerce mode, e.g. Taobao). This variable integrates corporate earnings report and expert interviews and is deduced from the iResearch statistical model. These data are continuous and unified.

National Bureau of Statistics released both monthly and annual data of TRSCG. To compare with seasonal data of online sales, we accumulated the monthly data between 2008 and June 2015 into seasonal data for correlation analysis. The TRSCG subtracted by online shopping market turnover was regarded as the proxy variable revealing the developing degree of online social commodity retail market, marked as OFS.

As showed in Figs. 1 and 2, the time series of ONS and OFS both vary and are affected by the seasonal factor. Since 2012, along with the acceleration of Internet development, the changing trends of online market sales and TRSCG are increasingly similar. To eliminate the effects of seasonal factor on these two series, we used Census X12 for seasonal adjustment. Also to remove the possible heteroskedasticity, we sent the adjusted series to logarithm, forming two series LONS_SA and LOFS_SA, respectively (Figs. 3 and 4).

![Figure 1. ONS of online TRSCG](image1)

![Figure 2. OFS of offline TRSCG](image2)

![Figure 3. Time series LONS_SA](image3)

![Figure 4. Time series LOFS_SA](image4)
3.2. Unit Root Test

As showed in Figs. 3 and 4, the seasonal factors in the original time series were eliminated, but the two series evidently tended to move to one direction with time, which was judged to contain an intercept item and a time item. Here Augmented Dickey-Fuller (ADF) unit root test containing constants and a time trend option was used to test the stability of series LONS_SA and LOFS_SA, which avoided the occurrence of spurious regression. The optimal lag order number was determined as per Akaike information criterion (AIC), where a smaller AIC means a better lag order number. The testing results are listed in Table 1.

Table 1 Unit root test of time series

<table>
<thead>
<tr>
<th>Variable</th>
<th>Test type (c,t,n)</th>
<th>ADF value</th>
<th>1% threshold</th>
<th>5% threshold</th>
<th>P value</th>
<th>Stationarity test</th>
</tr>
</thead>
<tbody>
<tr>
<td>LONS_SA</td>
<td>c,t,0</td>
<td>-1.9651</td>
<td>-4.3743</td>
<td>-3.6032</td>
<td>0.5912</td>
<td>Nonstationary</td>
</tr>
<tr>
<td>LOFS_SA</td>
<td>c,t,0</td>
<td>-1.1215</td>
<td>-4.3393</td>
<td>-3.5876</td>
<td>0.9064</td>
<td>Nonstationary</td>
</tr>
<tr>
<td>DLONS_SA</td>
<td>c,t,0</td>
<td>-6.2141</td>
<td>-4.3743</td>
<td>-3.6032</td>
<td>0.0002</td>
<td>Stationary</td>
</tr>
<tr>
<td>DLOFS_SA</td>
<td>c,t,0</td>
<td>-4.4754</td>
<td>-4.3240</td>
<td>-3.5806</td>
<td>0.0071</td>
<td>Stationary</td>
</tr>
</tbody>
</table>

Note: In the test type (c,t,n), c stands for the presence of an intercept item (0 means absence of this item); t for the presence of a time trend item (0 means absence of this item); s for a lag order number.

As showed in Table 1, the ADF statistics of LONS_SA and LOFS_SA both exceed the thresholds at significant levels 1% and 5%, indicating the null hypothesis cannot be rejected, the series are non-stationary and there unit roots. However, the ADF statistics of their first-order differential series DLONS_SA and DLOFS_SA are both smaller than the 1% threshold, indicating these two series are steady. In conclusion, LONS_SA and LOFS_SA after first-order differential smoothing both become first-order integrated variables, or namely I(1).

4. EMPIRICAL STUDY ON EFFECT OF ONLINE SHOPPING DEVELOPMENT ON TRSCG

4.1. Principles of Cointegration Test and VEC

4.1.1. Cointegration test

For a k-dimension vector time series \( y_t = (y_{1t}, y_{2t}, ..., y_{kt})^\prime \) (\( t = 1, 2, ..., T \)), let its component series be d,b-order cointegration, marked as \( y_t \sim CI(d,b) \). If \( y_t \sim I(d) \) is satisfied, then every component of \( y_t \) is d-order single integral; (2) There is a nonzero vector \( \beta \) that makes \( \beta' y_t \sim I(d-b) \), \( 0 < b \leq d \). For short, \( y_t \) is cointegrating, and vector \( \beta \) is called a cointegration vector (James, 2012).

In the model VAR(p), we suppose variables \( y_{1t}, y_{2t}, ..., y_{kt} \) are all non-stationary first-order single integer series, or namely \( y_{t} \sim I(1) \). Moreover, \( x_t \) is a d-dimension exogenous vector, representing the trend item and constant term.

\[
y_t = A_0 y_{t-1} + ... + A_p y_{t-p} + B x_t + \varepsilon_t, \quad t=1,2,...,T \quad (1)
\]

where \( y_{1t}, y_{2t}, ..., y_{kt} \) are nonstationary first-order integrated process \( I(1) \); \( X_t \) is a d-dimension exogenous disturbance vector. Then \( y_t-1 \) was subtracted from both sides, and through this addition and subtraction, then

\[
\Delta y_t = \sum_{i=1}^{p} \Gamma_i \Delta y_{t-i} + \Pi y_{t-i} + B x_t + \varepsilon_t \quad (2)
\]

\( I(1) \) was differentially transformed to a steady zero-order integrated process \( I(0) \). In Eq. (2), \( \Delta y_t, \Delta y_{t-j} (j=1,2,...,p) \) is a vector composed by \( I(0) \). Then as long as \( \Pi y_{t-i} \) is a \( I(0) \) vector, or namely there is cointegration relationship among the components of \( y_{t-1} \), then it is ensured \( \Delta y_t \) is a steady process.

A test regression coefficient method based on VAR proposed in 1988 by Johansen is feasible for multivariate cointegration test. The results of unit root test indicate all indices are \( I(1) \) series. The binary Johansen cointegration test (Table 3) show that the online and offline TRSCGs are cointegrative. On this basis, we further built a VEC model.

4.1.2. VEC model

Engle and Granger built a new VEC model by combining the cointegration and error correction model (ECM). In the VAR model, each equation stands for a self-regression distribution lagged model. Thus, it is believed that the VEC model is a VAR model under cointegration restraint and is mostly applied to build cointegration non-stationary time series models.

The VEC model is expressed as follows:
\[ \Delta Y_t = \alpha ECM_{t-1} + A_1 \Delta Y_{t-1} + A_2 \Delta Y_{t-2} + \cdots + A_p \Delta Y_{t-p} + \varepsilon_t \]  

(3)

where ECM stands for the error correction item computed from the cointegration equation. The error correction item reflects the non-equilibrium errors that the variables deviate from long-term equilibrium relations.

The coefficient \( \alpha \) before the error correction item is an adjustment parameter and used to reveal the speed how a variable is regressed from the current level to long-term equilibrium relation, or how the non-equilibrium errors are eliminated (Ruey, 2012).

The error correction item computed from this equation is exactly the variable CointEQ1 in ECM.

4.2. Modeling of JJ Cointegration Test and VEC

In the cointegration test, was used to regress and other appointed exogenous variables, and the lag order number of VEC means the first-order differential lag. Thus, prior lag order number determined from this principle.

<table>
<thead>
<tr>
<th>Lag</th>
<th>LR</th>
<th>FPE</th>
<th>AIC</th>
<th>SC</th>
<th>HQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>147.4671*</td>
<td>3.47e-07</td>
<td>-9.202035</td>
<td>-9.1239</td>
<td></td>
</tr>
</tbody>
</table>

Note: * means the optimal lag order number determined from this principle.

As showed in Table 2, the optimal lag order number is 1 or 5. Since the lag length of VEC is the lag length of the unconstrained VAR subtracted by 1, so the lag order number of VEC here would be 0, which is insufficient to affect the dynamics of the variables. Thus, here we selected the lag order number to be 5. The JJ cointegration test results of online and offline TRSCGs (Table 3) indicate the two variables are cointegrative.

<table>
<thead>
<tr>
<th>Hypothesized</th>
<th>No. of CE(s)</th>
<th>Eigenvalue</th>
<th>Statistic</th>
<th>Critical Value</th>
<th>Prob.**</th>
</tr>
</thead>
<tbody>
<tr>
<td>None *</td>
<td>0.612105</td>
<td>31.87275</td>
<td>20.26184</td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>At most 1</td>
<td>0.279557</td>
<td>8.197239</td>
<td>9.164546</td>
<td>0.0761</td>
<td></td>
</tr>
</tbody>
</table>

Table 3 Results of JJ cointegration test

Table 4 shows the regression results of vector error model. The first part of regression results in Table reveal the long-term equilibrium relation between two variables and the cointegration correlation is expressed as:

\[
\text{Coint EQ1} = \text{LOFS\_SA} - 0.092402 \text{ LONS\_SA} - 10.63279
\]  

(4)

Eq. (4) indicates that in the long run, when the online TRSCG increased by 1% and other conditions were unchanged, the offline TRSCG would rise by 0.092402%.

<table>
<thead>
<tr>
<th>Cointegrating Eq:</th>
<th>CointEq1</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOFS_SA(-1)</td>
<td>1.000000</td>
</tr>
<tr>
<td>LONS_SA(-1)</td>
<td>-0.092402***</td>
</tr>
<tr>
<td>C</td>
<td>-10.63279**</td>
</tr>
</tbody>
</table>

Table 4 Regression results of VEC

(a) Part 1

<table>
<thead>
<tr>
<th>Error Correction:</th>
<th>D(LOFS_SA)</th>
<th>D(LONS_SA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CointEq1</td>
<td>-0.017193</td>
<td>-0.631536***</td>
</tr>
</tbody>
</table>
The part 2 in Table 4 reveals the short-term relationship between two variables. When other variables were unchanged, the phase-t variation of offline TRSCG could eliminate 1.193% of non-equilibrium error from the previous phase. When other variables were unchanged, the phase-t variation of online TRSCG could eliminate 63.1536% of non-equilibrium error from the previous phase.

Part 3 in Table 4 reveals the overall fitting of the model. The indices R², adjustable R², AIC and SC indicate the overall fitting effect of the model is good.

### 4.3. Predictions

#### 4.3.1. Static prediction

Figures 5 and 6 show the static prediction images of the model.

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Part 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R-squared</td>
</tr>
<tr>
<td></td>
<td>Adj. R-squared</td>
</tr>
<tr>
<td></td>
<td>Sum sq. resid</td>
</tr>
<tr>
<td></td>
<td>S.E. equation</td>
</tr>
<tr>
<td></td>
<td>F-statistic</td>
</tr>
<tr>
<td></td>
<td>Log likelihood</td>
</tr>
<tr>
<td></td>
<td>Akaike AIC</td>
</tr>
<tr>
<td></td>
<td>Schwarz SC</td>
</tr>
<tr>
<td></td>
<td>Mean dependnet</td>
</tr>
<tr>
<td></td>
<td>S.D. dependent</td>
</tr>
<tr>
<td>Determinant resid covariance (dof adj.)</td>
<td>1.51E-07</td>
</tr>
<tr>
<td>Determinant resid covariance</td>
<td>6.17E-08</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>136.5557</td>
</tr>
<tr>
<td>Akaike information criterion</td>
<td>-9.244457</td>
</tr>
<tr>
<td>Schwarz criterion</td>
<td>-8.220601</td>
</tr>
</tbody>
</table>

Note: The square brackets in Part 1 and Part 2 mean the t-test. ****: p<0.01; ***: p<0.05; *: p<0.1
4.3.2. Dynamic prediction

Figures 7 and 8 show the dynamic prediction images six phases ahead after sample extrapolation.

Figure 7. Dynamic prediction image of LONS_SA  Figure 8. Dynamic prediction image of LOFS_SA

The prediction images show that the fitting effects of static prediction and dynamic prediction are both good, which further validate the reasonability and effectiveness of VEC.

5. CONCLUSIONS AND SUGGESTIONS

5.1. A Dynamic Coupling Effect Between Online and Offline Social Commodity Retail Markets was Found

Online retail sales tended to gradually invade offline retail sales from the perspective of retail sales of customer goods. Nevertheless, through fitting and prediction, we found a significant dynamic coupling effect between online and offline retail sales. In the long run, when the online TRSCG of LONS_SA increased by 1% and other conditions were unchanged, the offline TRSCG of LOFS_SA would rise by 0.092402%.

The error correction model reveals a significant short-term dynamic adjustment mechanism between the online and offline TRSCGs. Due to the presence of an error correction item, a long-term equilibrium relation between them was achieved. When short-term fluctuation deviated from long-term equilibrium and other variables were unchanged, the phase-t variation of offline TRSCG could eliminate 1.193% of non-equilibrium error from the previous phase. When other variables were unchanged, the phase-t variation of online TRSCG could eliminate 63.1536% of non-equilibrium error from the previous phase.

5.2. The Integration of Online Sales And Traditional Offline Retail would Largely Promote Economic Development

The integrated development mode combining offline retailing and online sales (O2O) has been increasingly applied by enterprises, especially Gome and Suning, which makes the “O2O” mode become a trend under the waves of “Internet+”. Compared with the traditional retailing, online shopping is more convenient, rapid and cheaper, but security and quality problems deserve attention. Under the physical shop sale mode in traditional retailing, customers see the physical commodities and are guaranteed with product quality, but the range of selection is narrower than online shopping. The online and offline integrated development solves the limitations of this single mode and more caters to the shopping habits of modern customers. The mutual promotion, complementation and coordinated development between the offline and offline shopping modes would
efficiently accelerate the development of social commodity retailing and largely urge the healthy growth of national economy.

5.3. **Internet Supervision and Online Enterprise Supervision**

As the online sales mode becomes increasingly popular, the risks of Internet frauds should never be ignored, given the virtuality of the online shopping environment. State supervision and management departments should study and strengthen the Internet environment surveillance and management and persistently promote the real-name registration, supervision and management of online enterprises/individuals, which guarantee the security of network transaction. Only when the Internet environment security is guaranteed, we are able to efficiently solve the safety concerns of online shopping and accelerate the circulation of consumption and commodities, thereby increasing the persistent growth of social commodity retail.

In conclusion, there is a strong dynamic coupling effect between online and offline retails, which tend to develop in a mutual promoting and joint dynamic trend. The online and offline shopping integration is deemed to be a trend in the social commodity retail market. As Internet supervision and management will be improved, the availability and accuracy of Internet monitored data would be enhanced, which help to more accurately analyze and predict the social commodity sales. Nevertheless, the hidden risks of online shopping cannot be ignored, which urge the government and relevant departments to take urgent measures that promote the rapid healthy development of online shopping and promote the long-term development of retailing in China.

**ACKNOWLEDGMENTS**

This work was supported by *Rural Finance Service Key Technology and System Development & Demonstration* from China National Science & Technology Pillar Program during 12th Five-Year Plan(2014BAL07B03); and Central University Basic Research Exclusive Fund from Beijing Language & Culture University Institute-Level Research Program (15YJ0408).

**REFERENCES**


