Forecasting Chinese Visitors to Japan
An Application of Cointegration and the Vector Error Correction Model

Kazutaka Kurasawa

Abstract
This study statistically analyzes the Chinese travel demand to Japan. The vector error correction model (VECM) of the three variables – the number of Chinese visitors, their real GDP per person and travel costs – is estimated by Johansen method, and the cointegrating vectors are identified. The dynamic relationships are described by the estimated vector autoregression (VAR), using the impulse responses and the variance decompositions. The forecasts are also made for the forecasting horizons of 2016 to 2020.

Keywords
Vector Error Correction Model, Travel Demand, Cointegration

1. Introduction
In recent years, Japanese travel industry has experienced an unprecedented boom in foreign visitors. This boom has been largely driven by the visitors from mainland China. In 2015, more than 4.4 million Chinese arrived in Japan, and mainland China was the largest origin country of foreign visitors, followed by South Korea (4.3 million), Taiwan (3.6 million) and Hong Kong (1.5 million), out of the total of 19.7 million visitors (the Ministry of Justice of Japan (2016)). From 1985 to 2000, the number of arrivals from mainland China to Japan expanded at an average rate of more than 10 percent per annum from around 100 thousand arrivals in the 1980s, surpassing 1.7 million in 2010 (Figure 5 plots the historical data of Chinese arrivals with their forecasts made in Section 4). Although it dropped to about 1.3 million in the aftermath of the earthquake and tsunami in 2011, Chinese travelers have returned to Japan with 2012, 2014 and 2015 breaking all-time records.

There are many potential driving forces behind the recent surge of Chinese visitors to Japan, such as political liberalization and globalization. Economic growth is also one of the drivers that have led many Chinese to travel overseas. A large increase in disposable income over the past decades have expanded opportunities for them to choose foreign countries as destinations (see, for example, Arlt and Burns (2013) for the recent development of Chinese international tourism). From the early 1990s to 2015, Chinese real GDP per capita increased roughly a sevenfold (Figure 2). During this period of time, the number of departures from China for international tourism increased more than 20 times (World Tourism Organization (2015)).

Assuming that the income elasticity of Chinese demand for international tourism is positive, the tourism demand from China to Japan will continue to increase in the foreseeable future since Chinese economy is expected to grow at a moderate, if not spectacular,
rate. International Monetary Fund (2016), for example, forecasts that Chinese Gross Domestic Product will increase by about 40% in real term from 2016 to 2020. It is, however, uncertain whether the rapid increase of the travel demand observed in the last years will be sustainable. There are some factors that seem to have helped temporarily boost Chinese tourism in Japan. For example, the Japanese government has relaxed visa rules for several countries, including China, to achieve the goal of 20 million foreign visitors by the year of the Tokyo 2020 Olympics. A weaker yen might also have lowered traveling and shopping costs in Japan. Finally, the recent surge of Chinese visitors might have been just a recovery to an underlying trend from the sharp drop after the 2011 earthquake. It is, therefore, not unlikely that the growth rate of Chinese travelers to Japan will decelerate to a long-run path after these temporary positive effects dissipate.

This study statistically analyzes and forecasts the Chinese travel demand to Japan from both short-term and long-term perspectives, applying the vector error correction model (VECM). The dynamic relationships between three variables - the number of Chinese visitors to Japan, their income and the travel cost – are considered in the VECM. The VECM generally describes the short-run dynamics of variables that have equilibrium relationships in the long run. Many times series variables individually wander apart, but some of them are bounded by equilibrium relationships. These variables are said to be cointegrated. Although cointegrating variables temporarily drift away from one another, their equilibrium relationships will be restored over time. The VECM is used to describe how cointegrating variables deviate from equilibrium in the short run and return to it in the long run. Empirically, it should be statistically tested whether variables are cointegrated or not. In addition, there can be more than one cointegrating relationship between a set of variable. Johansen (1991) develops statistical tests for cointegration in the VECM to determine the number of cointegrating relationships. This study applied these tests to the three variables.

In the literature of tourism management, demand for international tourism has been econometrically studied, in particular, since the 1990s, when the expansion of international travels has stimulated academic and practical interests in understanding the determinants of international tourism demand (see Li et al (2005) and Song and Li (2008) for the survey). For businesses and policy makers, accurate forecasts of expected demands are crucial since they make decisions, such as recruitment and capital expenditures, on the basis of long-term and short-term perspectives. Econometric approaches to modeling and forecasting Chinese visitors are, therefore, of considerable importance and interest to Japanese travel industry. With the unprecedented increase in Chinese travel demand, it is highly probable that simple extrapolation from the recent data will generate large forecast errors. The VECM adequately identifies the long-term and stable demand relationships and simultaneously estimate short-run deviations from equilibrium.

The rest of the paper is organized as follows. The next section describes the data used in this study and tests the stationarity of the variables. Stationarity is a prerequisite for cointegration analysis. The augmented Dickey-Fuller test for unit root is performed on the three variables. Section 3 presents the VECM. Section 4 reports the empirical results. The last section concludes.

2. Data and Unit Root Tests
2.1 Data

This study analyzes the relationships among the number of Chinese visitors to Japan, their income and the travel cost, using
annual time series for the period 1985-2013. The tourist arrival data is obtained from the Ministry of Justice of Japan, which categorizes Chinese data by region: "China", "China (Taiwan)", "China (Hong Kong)", and "China (Others)". In this study, "China" is used to count the number of arrivals from mainland China. The number of arrivals is divided by the size of population. The data of the population is also from the World Economic Outlook Database. For real GDP per capita, the data is sourced from IMF’s World Economic Outlook Database October 2015. Figure 1 plots Chinese real GDP per capita and population from 1985 to 2015.

For the cost of travel, the relative price level between China and Japan is calculated from the price indexes and the foreign exchange rate. In microeconomic theory, it is generally assumed that demand is a function of its own price. In practice, however, it is difficult to directly measure travel costs, such as transportation and living costs, from official data. In empirical studies, the price level of a destination country relative to an origin country is instead used as a proxy under the assumption that the relative price influences potential travelers’ decisions on where they travel (Song et al. (2009)).

The relative price index is defined as

\[
\frac{P_j/E_j}{P_c/E_c} \quad (1)
\]

where \( P_j \) and \( P_c \) are the GDP deflators in Japan and China, and \( E_j \) and \( E_c \) are the foreign exchange rates expressed in their national currencies per US dollar. The data of the GDP deflators are sourced from IMF’s Economic Outlook Database October 2015. The data of the exchange rate are obtained from FRED of the Federal Reserve Bank of St. Louis. The index is normalized as 1 in 1985. Figure 2 displays the normalized relative price over the sample period.

2.2 Unit Root Tests

Cointegration analysis requires that vari-
ables in a cointegrating system are of the same order of integration. Thus, the order of integration of the variables must be determined by unit root test before estimating cointegrating relationship (Engle and Granger (1991); Luo et al., (2007)). A variable is integrated of order 1 if its first difference does not have a unit root while its original series does (higher orders of integration are not considered in this study). Dickey and Fuller (1979) develops a statistical test for unit roots. This paper carries out the augmented Dickey-Fuller (ADF) test on the three variables: the number of arrivals from China to Japan per million persons \(y_{1t}\), Chinese real GDP per capita \(y_{2t}\) and the relative price index \(y_{3t}\). For \(y_{it}\), the ADF test is based on the auxiliary regression

\[
\Delta y_{it} = \alpha_0 + \alpha_1 t + \alpha_2 y_{it} + \sum_{j=1}^{l} \phi_j \Delta y_{it-j} + u_{it} \tag{2}
\]

where \(t\) is a deterministic time trend, \(u_{it}\) is an error term, and \(\alpha_0, \alpha_1, \alpha_2\) and \(\phi_1\) are all fixed coefficients. The lagged values of \(\Delta y_{it-j}\) are included to correct autocorrelation in \(u_{it}\). The lag length \(l\) is determined by information criterions, such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) (Akaike (1973) and Schwarz (1978)). The regression includes the time trend \(t\) and the drift term \(\alpha_0\), not to reduce the power of the test if it is statistically significant.

The null and alternative hypotheses are

\[
H_0 : \alpha_2 = 0 \\
H_1 : \alpha_2 < 0 \tag{3}
\]

The null hypothesis is that the variable has a unit root and is non-stationary. The alternative hypothesis implies that the variable does not have a unit root and is stationary. If the null hypothesis is rejected for the level of the variable, the test is performed for the first difference \(\Delta y_{it} = y_{it} - y_{it-1}\).

Table 1 presents the results from the ADF tests on the levels and the first differences. The p-values of the ADF statistics and the lag length selected by the information criterions are also reported in the table. In all cases, the AIC and the BIC select identical lag lengths.
For the levels, the time trends are all statistically significant. The null hypotheses are not rejected at the 5% significance level. For the first difference, the time trends are not statistically significant, but the drift terms are significant for $y_{1t}$ and $y_{2t}$ so that the ADF statistics reported in the table are those from the auxiliary regression without the deterministic time trend. For $y_{3t}$, the time trend is statistically significant. The null hypotheses are all rejected, which concludes that the variables are all integrated of order 1, or I(1).

3. Johansen’ Approach to Cointegration

Given that the variables are all integrated of the same order, they are to be tested for cointegration in order to see if there is a stable linear relationship between them in the long run. Roughly speaking, variables are cointegrated if non-stationary variables of the same degree of integration move together over time. Although cointegrating variables drift away from one another in the short run, their equilibrium relationship will be restored in the long run. A set of cointegrating variables can have multiple cointegrating relationships, and the number of cointegrating relationships $r$ is $0 \leq r < n-1$, where $n$ is the number of variables. Empirically, the number of cointegrating relationships must be determined by statistical tests.

Johansen (1998) develops the maximum likelihood technique to determine the number of cointegrating relationship. Consider a vector $y_t = [y_{1t}, y_{2t}, y_{3t}]'$. Johansen’s approach first sets up a VAR as

$$y_t = \sum_{j=1}^{l} B_j y_{t-j} + \epsilon_t$$

(4)

where $l$ is the lag length of the VAR, $B_j$ is a matrix of fixed coefficients, and $\epsilon_t$ is a vector of independent and identically distributed error terms.

Under the assumption that the variables are cointegrated, the VAR is turned into the VECM:

$$\Delta y_t = \sum_{j=1}^{l-1} \Gamma_j \Delta y_{t-j} + \Gamma y_t + \epsilon_t$$

(5)

where $\Gamma_j = (\sum_{j=1}^{l} B_j) - I$ and $\Gamma_j = (\sum_{j=1}^{l-1} B_j) - I$ (I is an identity matrix). The VECM describes what proportion of the disequilibrium is corrected from t-1 to t (Engle and Granger (1987)).

The focal point of Johansen’s approach is to estimate and test the matrix $\Gamma$. In the long-run equilibrium, $\Delta y_t$ are all zero for $j = 1, \cdots, l-1$. If $\epsilon_t$ is also set to zero, $\Gamma y_t = 0$. Thus, the matrix $\Gamma$ is interpreted as a matrix of long-run coefficients. $\Gamma_j$ is, on the other hand, interpreted as a matrix of short-run adjustments.

The matrix $\Gamma$ is decomposed as

$$\Gamma = a \beta'$$

(6)

$a$ is a $n \times r$ matrix of adjustment parame-

<table>
<thead>
<tr>
<th></th>
<th>$y_{1t}$</th>
<th>$y_{2t}$</th>
<th>$y_{3t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>level</strong></td>
<td>ADF</td>
<td>p-value</td>
<td>lag</td>
</tr>
<tr>
<td></td>
<td>-2.556</td>
<td>0.300</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>-3.536</td>
<td>0.055</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-1.776</td>
<td>0.690</td>
<td>0</td>
</tr>
<tr>
<td><strong>1st difference</strong></td>
<td>ADF</td>
<td>p-value</td>
<td>lag</td>
</tr>
<tr>
<td></td>
<td>-9.746</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>-4.226</td>
<td>0.013</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-5.478</td>
<td>0.001</td>
<td>1</td>
</tr>
</tbody>
</table>
Forecasting Chinese Visitors to Japan
(Kazutaka Kurasawa)

ters (or weighting elements), which represents the speed of adjustment to disequilibrium. \( \beta \) is a \( r \times n \) matrix of cointegrating vectors. In equilibrium, the variables are linearly bounded by the column vector of \( \beta \). The elements of \( \beta \) are, thus, interpreted as long-run elasticity if the data is log-transformed. The number of cointegrating vectors \( r \) can be determined by checking the significance of the eigenvalues of the matrix \( \Gamma \).

Assuming \( e_t \sim N(0, \Sigma) \), the matrix \( \Gamma \) is estimated by maximizing the log-likelihood function:

\[
-\frac{nT}{2} \log(2\pi) - \frac{T}{2} \log(|\Sigma|) - \frac{1}{2} \sum_{t=1}^{T} e_t' \Sigma^{-1} e_t,
\]

where \( T \) is the number of observations.

Johansen’s approach determines the number of cointegrating vectors \( r \) by testing whether the eigenvalues of the matrix \( \Gamma \) are significantly different from zero. The rank of a matrix is generally equal to the number of its eigenvalues that are different from zero. Thus, if there is no cointegrating relationships between variables in a VAR, the eigenvalues must be all zero.

There are two statistical tests for cointegration applicable to Johansen’s approach. One is called the trace test. The null hypothesis is that there are at most \( r \) cointegrating vectors against the alternative that there are more than \( r \) cointegrating vectors. The trace test is based on the trace of the estimated matrix \( \hat{\Gamma} \):

\[
\hat{\lambda}_{\text{trace}}(r) = -\frac{T}{2} \sum_{k=r+1}^{\infty} \ln(1 - \hat{\lambda}_k)
\]

where \( \hat{\lambda}_k \) is the \( i \)th estimated value of the eigenvalues in ascending order. The intuition behind the trace test is that the larger is the estimated eigenvalue, the larger is the test statistic, which rejects the null hypothesis.

The other test is known as the maximal eigenvalue test based on the test statistic:

\[
\hat{\lambda}_{\text{max}}(r+1) = -T \ln(1 - \hat{\lambda}_{r+1})
\]

This statistic tests the null hypothesis that there are \( r \) cointegrating vectors against the alternative of \( r + 1 \). The larger is the estimated eigenvalue \( \hat{\lambda}_{r+1} \), the more likely the test statistic is to reject the null hypothesis of \( r \).

The distributions of \( \hat{\lambda}_{\text{trace}} \) and \( \hat{\lambda}_{\text{max}} \) are not standard; Johansen and Juselius (1990) and Osterwald-Lenum (1992) provide critical values for these statistics.

The estimated VECM is transformed back into the VAR to investigate the dynamic interactions between the variables in the short run. The estimated VAR is used to compute impulse responses and variance decompositions. Impulse responses trace out the impact of a shock to one variable on the variables in the VAR, including itself, over time. Variance decompositions describes the proportions of the variations in one variable that are attributed to its own shock and shocks to the other variables. The estimated VAR is also used to make the n-ahead forecasts of the variables.

4. Empirical Results

In order to specify the VAR, the lag length for the VAR is selected by the AIC and the BIC. Table 2 reports the values of the AIC and BIC with \( l \) lag length \( l \) from 1 to 4. The values of the AIC and the BIC are both minimized for \( l = 3 \). The VECM with \( l = 3 \) is therefore estimated from the data (a constant term is also added to the cointegrating relationships).

Table 3 presents the test statistics \( \hat{\lambda}_{\text{trace}} \) and \( \hat{\lambda}_{\text{max}} \).
and $\lambda_{\text{max}}$ from these eigenvalues. The eigenvalues computed from the estimated matrix $\hat{C}$ are 0.8591, 0.5069 and 0.3601 in ascending order. At the 5% significance level, the tests reject all the null hypotheses, which concludes that there are two cointegrating vectors ($r = 2$).

The cointegrating vectors are presented in Table 4. The values in the table are normalized so that they can be interpreted as the long-run elasticities of Chinese visitors to Japan (the data are all log-transformed). The values of the second cointegrating vector all have the expected signs. The long-run elasticity of Chinese visitors to Japan is 1.178 with respect to Chinese real GDP per capita. The elasticity with respect to the relative price level is -0.351. In the first cointegrating vector, however, the real GDP per capita does not have the expected positive sign. Although the two vectors are identified from the VECM, only one of the cointegrating relationships is economically meaningful.

Given these cointegrating relationships, the VECM can be rewritten with the cointegrating residuals computed from the cointegrating vectors:

\[
u_{1t} = y_{1t} - 0.142y_{2t} - 0.819y_{3t} - 4.923
\]

\[
u_{2t} = y_{1t} + 1.178y_{2t} - 0.351y_{3t} - 8.303
\]

For each variable $y_{it}$,

\[
\Delta y_{it} = \sum_{j=1}^{t} \Gamma_{j} \Delta y_{i,t-j} + \gamma_{i1}u_{1t} + \gamma_{i2}u_{2t} + e_{i}
\]

where $\gamma$’s are fixed coefficients. The equations (5)’ are estimated by the ordinary least squares (OLS). Table 5 reports the OLS estimates and standard errors for this form of the VECM.

The estimated VECM is transformed back in the VAR. Table 6 reports the estimated matrix of $B_r$. With these estimates, the impulse responses of Chinese visitors to Japan are calculated. Figure 4 display the impulse responses of Chinese visitors to shocks to the three variables in the VAR. In computing the impulse responses, recursive structure is imposed on the variance covariance matrix of $e_t$ so that shocks to real GDP per capita ($y_{2t}$) and the travel cost ($y_{3t}$) are assumed not to contemporaneously affect Chinese travel demand to Japan ($y_{1t}$). The size of the shocks is one standard deviation. The figure shows that the number of Chinese visitors continuously increase over 5 years after a positive shock hits real GDP. A shock to the travel cost, on the other hand, have lagged effects on the travel demand; the number of Chinese visitors to increase two years after the travel cost unexpectedly falls, but gradually returns to the equilibrium thereafter. The impulse responses indicate that real income has long-lasting effects on the travel demand while the travel cost does not. This is also shown con-
Table 5: The Estimates of the VECM

<table>
<thead>
<tr>
<th></th>
<th>$y_{1t}$</th>
<th></th>
<th>$y_{2t}$</th>
<th></th>
<th>$y_{3t}$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>estimate</td>
<td>s.e.</td>
<td>estimate</td>
<td>s.e.</td>
<td>estimate</td>
<td>s.e.</td>
</tr>
<tr>
<td>$u_{1t}$</td>
<td>-0.346</td>
<td>0.034</td>
<td>-0.298</td>
<td>0.351</td>
<td>-0.933</td>
<td>0.669</td>
</tr>
<tr>
<td>$u_{2t}$</td>
<td>0.035</td>
<td>0.009</td>
<td>-0.227</td>
<td>0.094</td>
<td>-0.219</td>
<td>0.179</td>
</tr>
<tr>
<td>$\Delta y_{1t}$</td>
<td>0.056</td>
<td>0.017</td>
<td>-0.086</td>
<td>0.170</td>
<td>-0.278</td>
<td>0.324</td>
</tr>
<tr>
<td>$\Delta y_{2t}$</td>
<td>0.265</td>
<td>0.142</td>
<td>-0.829</td>
<td>1.465</td>
<td>1.522</td>
<td>2.791</td>
</tr>
<tr>
<td>$\Delta y_{3t}$</td>
<td>-0.015</td>
<td>0.019</td>
<td>0.182</td>
<td>0.197</td>
<td>-0.484</td>
<td>0.375</td>
</tr>
<tr>
<td>$\Delta y_{1t}$</td>
<td>0.177</td>
<td>0.022</td>
<td>0.420</td>
<td>0.231</td>
<td>0.242</td>
<td>0.441</td>
</tr>
<tr>
<td>$\Delta y_{2t}$</td>
<td>-1.043</td>
<td>0.120</td>
<td>-0.756</td>
<td>1.235</td>
<td>-3.648</td>
<td>2.353</td>
</tr>
<tr>
<td>$\Delta y_{3t}$</td>
<td>0.041</td>
<td>0.020</td>
<td>-0.401</td>
<td>0.204</td>
<td>-0.444</td>
<td>0.389</td>
</tr>
</tbody>
</table>

Table 6: The Estimated Matrixes for the VAR

<table>
<thead>
<tr>
<th></th>
<th>$\hat{B}_1$</th>
<th></th>
<th>$\hat{B}_2$</th>
<th></th>
<th>$\hat{B}_3$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$y_{1t}$</td>
<td>$y_{2t}$</td>
<td>$y_{3t}$</td>
<td>$y_{1t}$</td>
<td>$y_{2t}$</td>
<td>$y_{3t}$</td>
</tr>
<tr>
<td>$y_{1t}$</td>
<td>1.092</td>
<td>-0.131</td>
<td>0.832</td>
<td>-3.702</td>
<td>0.128</td>
<td>0.400</td>
</tr>
<tr>
<td>$y_{2t}$</td>
<td>1.248</td>
<td>-0.001</td>
<td>0.060</td>
<td>-1.250</td>
<td>0.059</td>
<td>0.117</td>
</tr>
<tr>
<td>$y_{3t}$</td>
<td>-1.158</td>
<td>1.451</td>
<td>-0.002</td>
<td>1.197</td>
<td>-0.515</td>
<td>0.413</td>
</tr>
</tbody>
</table>

Figure 3: The Impulse Responses of Chinese Visitors to Japan
firmed by the variance decomposition on Figure 5.

The number of Chinese visitors to Japan per million are forecasted from the VAR for the 5 year horizon from 2016 to 2020. The 95% confidence intervals are also computed by bootstrapping (1000 runs). The point forecasts and the confidence intervals are reported in Table 7. Figure 5 plots the forecasts of Chinese visitors to Japan in million persons from 2016 to 2020 with the 95% bootstrap forecast intervals. The forecasts are conditional on IMF’s estimates of the population. The figure also displays the historical data of Chinese visitors from 1985 to 2015. According to the point forecasts, about 4.7 million Chinese are expected to visit Japan in 2016, and 17.5 million in 2020. The uncertainty of these statistical forecasts is, however, not negligible, particularly, at longer forecast horizons. With a 95% probability, the number of Chinese visitors to Japan is forecasted to lie within 3.4-6.5 million in 2016, and 6.7-46.1 million in 2017.

6. Concluding Remarks
This paper has analyzed the demand of Chinese visitors to Japan using the VECM. The cointegrating vectors have been identified and the long-term elasticities have been estimated. The elasticities imply that the number of Chinese visitors per million in-

Table 7: The Forecasts of Chinese Visitors to Japan per Million Persons from 2015 to 2020

<table>
<thead>
<tr>
<th>Year</th>
<th>95% lower bound</th>
<th>Forecast</th>
<th>95% upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>2434.601</td>
<td>3387.085</td>
<td>4712.209</td>
</tr>
<tr>
<td>2017</td>
<td>3243.700</td>
<td>5021.309</td>
<td>7773.082</td>
</tr>
<tr>
<td>2018</td>
<td>4303.948</td>
<td>7656.184</td>
<td>13619.393</td>
</tr>
<tr>
<td>2019</td>
<td>4586.590</td>
<td>9893.784</td>
<td>21342.016</td>
</tr>
<tr>
<td>2020</td>
<td>4734.579</td>
<td>12441.302</td>
<td>32692.697</td>
</tr>
</tbody>
</table>
creases roughly 1.7% in equilibrium for every 1 percent growth in real GDP per capita, and it decreases 0.4% for every 1 percent increase in travel cost. The impulse responses and the variance decompositions computed from the estimated models indicate that the income has quantitatively large and long-lasting effects on the Chinese travel demand to Japan while travel costs do not. The forecasts have also been made from the estimated model. In 2020, for example, about 17.5 million Chinese are forecasted to visit Japan.

References


