Purchasing Power Parity Theory Determinants –A Swedish Destination Study of International Tourists: a Count Data Approach

A. Khalik Salman
Department of Business
Economics and Law
Mid-Sweden University
Sweden

Ghazi Shukur
Department of Economics and Statistics
Linnaeus University
Department of Economics and Statistics
Jönköping University
Sweden

Abstract

This paper employs the time-series negative binomial regression model (TNBM) to test the hypothesis effects of purchasing power parity (PPP) theory on the counts data of visitors to the north-west of Sweden (SW6 region). We consider a sample of monthly time-series count data from 1993:01 to 2008:12 taken from five countries: Denmark, the United Kingdom, Switzerland, Japan and the United States. For each visiting country, we specify separate equations by including the relative available information. We then estimate these equations using the time - series negative binomial model (TNBM). The benefit of this model is that it is much more flexible and therefore likely to fit better (if the data is not Poisson distributed) and hence is more efficient than single-equation estimation methods such as least squares. We found that the number of visitors to Sweden is negatively related to the absolute PPP and relative PPP. This result is in accordance with macroeconomic theory and the PPP theory. The results also show that some lagged dependent variables, and several monthly dummies (representing seasonal effects), have a significant impact on the number of visitors to north-west Sweden. We also find that, in at least some cases, absolute PPP, relative PPP and relative price have significant effects on international tourism demand.

Keywords: Tourism Demand, PPP Theory Approach, Time-Series Negative Binomial Model

1. Introduction

Aggregate demand theory constitutes a central topic in macroeconomic theory, and modern formalized macroeconomic theory has dealt with consumer demand for goods and services for some time. This has led to Macro econometrics applications of demand function and how they related to applied to international tourism demand is the theme of this paper.

During the past decade or so, the literature on tourism demand has included a number of different statistical evaluations used to identify the relationships between the number of tourists arriving in a particular country and the factors that influence these arrivals. In international tourism demand modeling, most studies have used a demand function approach to identify quantitative relationships. However, from a methodological point of view, separate models of visitor numbers can be estimated by equation count data such as negative binomial regression models.

Many external and internal factors influence tourism demand. In turn, tourism generates physical and international financial flows that have potentially strong economic and environmental impacts. Consequently, there is a broad group of stakeholders in tourism, arising from both the private and public sectors. From these has emerged a widely felt need for analysis of tourism in the wider context of the national account that is to say nationally and internationally comparable with measures of other economic activities.

Importantly, in existing econometric studies of tourism in Scandinavia (particularly Sweden), factors such as absolute purchasing power parity (APPP), relative purchasing power parity (RPPP) and relative price to measure cost competitiveness have not been important determinants in international tourism demand models, with relatively more emphasis on seasonal effects.

In the former types of tourism econometric studies, factors such as RPPP, relative price and APPP have not been considered as the most important determinants for the international tourism demand models (Salman, 2003; Salman, Bergmann-Weinberg and Shukur, 2007; Salman, et al. 2010; Salman, 2011). Special events and nature have been considered as the most important determinants for tourism demand models in previous Scandinavian studies (Hultkrantzand Olsson, 1997). Additionally, there have been extensive empirical studies on Scandinavian tourism demand models using various econometric test procedures: cointegration analysis (Salman, 2003; Salman, Bergmann Weinberg, and Shukur 2007; Kronenbery, 2013); ISUR models (Salman et.al. 2010; Salman, 2011), and linear-regression models (Hultkrantz, 1995; Hultkrantzand Olsson, 1997).

The aim of this paper is to estimate international tourism demand for Sweden from five countries: namely, Denmark, the United Kingdom (UK), Switzerland, Japan and the United States (US). For each visiting country, we specify a separate equation with the relative information included in each equation. Previous Scandinavian studies have not used negative binomial regression. Further, previous studies of Swedish tourism demand have not used the RPPP, APPP and the relative price to measure the cost competitiveness. Yet there are other factors that influence demand for tourism include climate, cultural values, natural attractions and government travel regulations, many of which are difficult to quantify.

The objective of this study is to analyze how exchange rates affect the number of visitors to Sweden, using the PPP hypothesis, and how seasonal (monthly) conditions influence the tourism demand function, using time-series negative binomial model (TNBM). On the other hand PPP should hold better when comparing goods with perfect substitute and relatively low transport costs. This the empirical study aims to contribute to the debate about PPP and investigate if the PPP hold in the long run between Sweden and Norway as well

The remainder of the paper is organized as follows. Section 2 gives the variables of the study and the data used. Section 3 presents the estimation and testing methodology. Section 4 provides the results. The paper concludes with a brief summary and conclusion in Section 5.

2. Data Description

In this paper we use time-series data for the five countries such as Denmark, the United Kingdom, Switzerland, Japan and the United States over the period 1993:01 to 2008:12. We have built from original sources many of the variables used in our tourism demand model, as no available time-series datasets had the characteristics necessary to pursue our objective. In order to determine the international macroeconomic theory approach drivers of the tourism demand function within Swedish regions, indices such as APPP, RPPP and relative price have been adopted.

For each year the variables have been built by the following calculations.

i. Absolute purchasing power parity (APPP) version can be stated exchange rate: $S = P(CPI) / P(CPI)^*$, where "S" is the exchange rate defined as domestic currency units per unit of foreign currency, "P" is the price of goods expressed in the domestic currency, and "P^* " is the price of an identical bundle of goods in foreign country expressed in terms of the foreign currency. According to APPP, arise in the home price level relative to the foreign price level will lead to a proportional depreciation of the home currency against the foreign currency. In our study, if the prices in Sweden rise while prices in the other countries remain at the same level, then – according to APPP – the Swedish currency will depreciate. This variable measures the real cost of living in relative terms for foreign countries and Sweden.

ii. The relative purchasing power parity (RPPP) version of exchange rate: the absolute version of PPP is, proponents of the theory generally acknowledge, unlikely to hold precisely because of the existence of transport costs, imperfect information and the distorting effect of tariff and non-tariff barriers to trade. Nonetheless, it is argued that a weaker form of PPP know as relative PPP can be expected to hold even in the presence of such distortions. Put simply, the relative version of PPP theory argues that the exchange rate adjust by the amount of the inflation differential between two economies.

Algebraically this is expressed as: $\%\Delta S – \%\Delta P(CPI) - \%\Delta P(CPI)^*$, where $\%\Delta S$ is the percentage change in the exchange rate, $\%\Delta P$ is the domestic inflation rate and $\%\Delta P^*$ is the foreign inflation rate. According to RPPP, if the inflation rate in Sweden is 10% whilst the inflation rate in the foreign country is 4%, the Swedish krona per unit of foreign currency should be expected to depreciate by approximately 6%. The RPPP reflects the cost of living and the opportunity cost. This variable measures the percentage cost of living in relative terms for the foreign country and Sweden, and the percentage substitute price for a foreign tourist.

The PPP theory variables APPP and RPPP were used to measure relative terms for the foreign country and Sweden and a substitute price for a foreign tourist. Tourists, who visit Scandinavian countries like Sweden and Norway, come primarily to enjoy the nature and the skiing. And the two countries have been competing with each other to attract more tourists. The tourists have the option of spending vacations in Sweden or in Norway. Both have similar climates and geography. Therefore, from the point of view of potential visitors, Norway is considered to be a competitive and substitute destination for Sweden, and the cost-of-living variable for the tourism demand model is defined as the absolute and percentage relative (APPP and RPPP) cost of living in Sweden compared with that of Norway. Moreover, the PPP hypothesis was embodied in the monetary theory of the balance-of-payments approach, and this resulted in an empirical formulation which expresses exchange-rate movements in terms of relative money supply.

iii. Additionally, the relative price rate is a relevant factor in determining the effect of cost competitiveness on tourism demand. The rationale behind incorporation of the relative price as a separate explanatory variable is that tourists may be more aware of the relative price than the specific cost of tourism at the destination. A question that arises is whether the exchange rate should be included in our model system as an explanatory variable together with the price variable. Measures of cost competitiveness differ between models in trade equations. Some modelers prefer to use relative price whereas others use relative labour cost or even relative total cost. The empirical evidence does not seem to offer convincing support for one alternative over the others but simulation properties of full models may be sensitive to the different specifications. For example, relative price may change as a result of change in the price mark-up, whereas relative labour costs will be invariant to such a change. One repeated concern about empirical estimates of cost-competitiveness elasticities is that they are “too low” and do not always satisfy the static Marshall-Lerner condition, (Whitley, 1994).

We can specify the price of tourism at the destination in a variety of ways. For instance, we can represent CPI in either absolute or relative terms. However, we consider the relative price as a measure of cost competitiveness between Sweden and Norway in this study. We define this as the ratio of the consumer price index (CPI) of the host country (CPI_{SW}) to the country of origin adjusted by the relative exchange rate (EX_{it}) to obtain a proxy for the real cost of living (Salman, 2003). Therefore, the real costs of tourism in Sweden and Norway are the relative CPIs given by:

$$Rp_{jt} = \frac{\dfrac{\mathrm{CPI}_{it}}{EX_{ijt}}}{\mathrm{CPI}_{jt}}, \qquad (1)$$

where, i is the host country (Sweden or Norway), j is the visiting (or foreign) country, and t is time. Rp_{it} is the relative price, consumer price indices (CPI) for country i in time t, CPI_{it} is the CPI for Sweden or Norway, $CPI_{jt}$is the CPI for the foreign country, and EX_{ijt} is the exchange rate between the Swedish krona/Norwegian krone and the foreign currency.

Iv. We use dummy variables from January to November as proxies for seasonal effects (December is the base category).

2.1 The Specification of the Model

The objective of this section is to analyze how the following macroeconomic theory, PPP hypotheses and seasonal (monthly) conditions influence the international tourism demand for Sweden.

$$Visitor_t = \beta_0 + \beta_1 Appp_t + \beta_2 Rppp_t + \beta_3 RP_{tij} + \theta Dummy_i + \beta_4 Visitors_{t-i} + u_t. \qquad (2)$$

where:

Visitors= dependent variable measured thecountdata (number of visitors to Sweden).
APPP = absolute purchasing power parity exchange rate (S = P (CPI) / P(CPI)*.
RPPP = relative purchasing power parity exchange rate (%ΔS– %ΔP (CPI) -%ΔP (CPI).
RP_{tij}= relative prices are defined as follows,

where:

relative price of tourism for Denmark $= \frac{CPI_{SW} / EX_{SKr/DKr}}{CPI_{NO} / EX_{NKr/DKr}}$. (3)

relative price of tourism for the UK $= \frac{CPI_{SW} / EX_{SKr/GBp}}{CPI_{NO} / EX_{NKr/GBp}}$. (4)

relative price of tourism for Switzerland $= \frac{CPI_{SW} / EX_{SKr/SWf}}{CPI_{NO} / EX_{NKr/SWf}}$. (5)

relative price of tourism for Japan $= \frac{CPI_{SW} / EX_{SKr/JPy}}{CPI_{NO} / EX_{NKr/JPy}}$. (6)

relative price of tourism for the US $= \frac{CPI_{SW} / EX_{SKr/USD}}{CPI_{NO} / EX_{NKr/USD}}$. (7)

Where:

CPI_{Sw}: CPI in Sweden (1998 = 100).
CPI_{No}: CPI in Norway (1998 = 100).
$EX_{SKr/DKr}$: an index of the Swedish krona per unit of Danish krone (1998 = 100).
$EX_{SKr/GBP}$: an index of the Swedish krona per unit of British pound (1998 = 100).
$EX_{SKr/SWf}$: an index of the Swedish krona per unit of Swiss franc (1998 = 100).
$EX_{SKr/JPY}$: an index of the Swedish krona per unit of Japanese yen (1998 = 100).
$EX_{SKr/USD}$:an index of the Swedish krona per unit of US dollar (1998 = 100).
Dummy: the monthly dummies variables as proxies for seasonal effects.
Visitors $_{t-i}$: lagged dependent variable.

In this paper, we attempt to explain international flows of tourists to the north-west of Sweden (SW6 region) and to the Tröndelag region in mid-Norway, which is an alternative destination to the objective (SW6)[1] from Denmark, the UK, Switzerland, Japan and the US. Therefore, we define the cost-competitiveness effects between these two regions by relative price (RP_{tij}) as the ratio of the CPI of the host country (CPI_{SW} and CPI_{NO}) to the country of origin (CPI) adjusted by the relative exchange rate. This provides a proxy for the real cost of living. We define the relative price in Sweden and Norway by relative CPI_{SW} and CPI_{NO} as follows, along with the cost competitiveness (in relative prices):

As for the signs of the explanatory variables, we expect a negative sign for the relative-price variable and a positive sign for the exchange-rate variable. A lagged dependent variable may also be included to account for habit persistence and supply constraints. In this study, monthly dummies represent seasonal effects on the number of arrivals from the origin countries. All the independent variables are in natural logarithms, and the data are in index form (1998 = 100). All economic data employed in this study are from Statistics Sweden (Statistiska Centralbyrån) and Statistics Norway. We used E-Views Ver. 8.1 statistical program packages for the estimation. We examine monthly time-series count data from 1993:01 to 2008:12.

The SW6 region is a major tourist destination worldwide, with the yearly tourism demand in this part of Sweden consistently following an upward trend (see the map 1 in appendix B of this study).

[1] In our case, tourists consider Tröndelag an alternative to the objective SW6 region. These are the two destinations in Scandinavia, at least in terms of arrivals, for tourist from the origin countries under consideration. (see Figures 1 in the appendix B of this study)

However, interruption to these trends has taken place on a number of occasions due to macroeconomic factors and PPP hypotheses having a detrimental effect on tourism demand for Sweden.

A common model used in tourism-demand studies is a single equation with demand explained by the tourists' income in their country of origin, the cost of tourism in their chosen and alternative destinations, and a substitute price (Salman, 2003; Salman et al. 2010; Salman, 2011). To start with, the tourism demand can be expressed in a variety of ways. The most appropriate variable to represent demand explained by economic factors is consumer expenditure or receipts (Salman, 2003). Other measures of demand are potentially the nights spent by tourists or their length of stay. However, due to the lack of data on monthly GDP, personal income (GDP/population) is not included in this analysis.

The tourism price index (the price of the holiday) is also an important determinant of the decision a potential tourist makes. We can divide this into two components: (i) the cost-of-living index for the tourist at the destination, and (ii) the cost of travel to the destination. We divide the cost of living into two components: (i) the APPP form, assuming that tourists have the option of spending their vacation in Sweden, and (ii) RPPP tourist consumer expenditure or real consumer expenditure.

3. Methodology

Many researchers have investigated tourism demand in the context of individual Scandinavian countries, but no one has researched it using a pan-county approach. We think that a pan-country approach may have some bias because of the heterogeneity of different countries. Monthly time-series count data from 1993:01 to 2008:12 are used. We selected this period because it is prior to the sub-prime crisis and the effects of this crisis on the world and European economies, which we take to start from the public's awareness of the bankruptcy of Lehman Brothers on the 15th September of 2008.

However, we decided to use time -series negative binomial model (TNBM) as the estimator for count data. In order to estimate the factors that influence the number of visitors to Sweden, we analyze the characteristics of the number of visitors' variable. The number of visitors is a count data; this kind of variable cannot be negative because we cannot have negative number of visitors. There can be no visitors but not minus numbers of tourists, so a negative value of this variable would be a nonsense. Also we have another constraint: the number of visitors is always an integer number, so a 'half-number' which would also be a nonsense. So we need an estimator that can be robust to these two constraints. An ordinary-least-squares (OLS) estimator can be used with log-transformation of count variable, for non-integer data, but it is also not possible to use this approach where the count data assumes the value of zero, because we cannot have log(0). Hence, we thought that models based on the classic OLS estimator was not appropriate, (Tabachnick and Fidell, 2007). Numerous techniques have been developed for count data such as Poisson, negative binomial, zero-inflated poison (zip) and zero-inflated negative binomial (Long and Freese,2006;Sano et al., 2005). These techniques can handle non-normality on the dependent variable and do not require the researcher to either dichotomize or transform the dependent variable. We focus on the negative binomial technique.

The negative binomial distribution is similar to Poisson distribution, but the assumption of independence of observations is lifted, reflecting the notion that the extent to which a participant engages in repeated occurrences may be influenced by individual differences (Sturman, 1999). Further, the variance and mean are not assumed to be equal, so over-dispersion is no longer problematic. These assumptions aside, the similarity between negative binomial and Poisson techniques are demonstrated by the fact the negative binomial distribution converges to Poisson distribution when the variance and mean are equal. The statistical studies confirmed that the negative binomial regression is much more flexible and therefore likely to fit better, where the data are not Poisson distributed (Hausman, Hall and Griliches, 1984).

Moreover, the empirical literature confirms that in most cases there are two ways to use count data without Poisson regression, either the quasi-maximum likelihood (QML) Poisson regression or the negative binomial regression (Verbeek, 2008). The QML Poisson regression still presents a problem, in our case because the sample is large (156 observations), so the easiest alternative is to use the negative binomial regression as described in Hausman (1984). To estimate the effects of the macroeconomic variables and PPP theory variables on the count data of visitors to Sweden we have decided to use the time- series negative binomial. The HHG estimation procedure, the most commonly used procedure in statistical software for fixed-effects NBE (FENBE), does not qualify as a true fixed-effects method, because it does not control for unchanging covariates.

In fact, as explained by Allison and Waterman (2002), the problem with the HHG-FENBE is that it allows for individual-specific variation in the dispersion parameter rather than in the conditional mean. So the time-invariant covariates can appear statistically significant when they are not.

The statistical results show that the main relationship between the number of visitors coming to Sweden and the majority of the independent variables is in accordance with PPP theory and macroeconomic theory. The estimation of the model is achieved by the use of time –series negative binomial model. The results show that the negative binomial regression result is valid and adopted in this study. Using the valid method we also show the conformity between the expected signs for coefficients and those obtained as theoretical criteria. Indeed, a decrease in home CPI (destination cost of living) and the existence of relative price should engender an increase in the number of visitors. Therefore, the signs of these variables (APPP, RPPP) should be negative. Whereas, an increase in relative price should result in a decrease in the number of visitors to Sweden, so we would expect that the sign of this variable would be negative. Therefore, we can also conclude that this finding links with PPP theory hypothesis and macroeconomic variables

4. Results

In this section we present and discuss the results in following order: first, determination of which model provides the best fit with the observed data; and second identification and interpretation of significant predictors. All regression results and analyses were conducted with the statistical software, E-Views version 8.1.

4.1 Estimation Results

This study employs monthly count data covering the period 1993:01 to 2008:12. In this section we present our most important results, using the TNBM to find out what kind of factors have an effect on the number of visitors to Sweden from five countries. First, we wrote a theoretical specification, which consisted of the five variables described above. In addition to the major macroeconomic factors, the model also included the PPP-theory approach variables, such as the APPP and RPPP factors, to get a satisfactory explanation for their effect on the dependent variable. To improve the robustness of the results, the sample starts from 1993:01 andends in 2008:12, excluding the shocking effects of the sub-prime crisis on the economic system. From the results presented in the Table 1, and by looking at the LR-test, which is a test of the over-dispersion parameter (alpha), it is clear that these results are affected by over-dispersion. When the over-dispersion parameter is zero the negative binomial distribution is equal to a Poisson distribution (Washington et al., 2003). Further, alpha is significantly different from zero, and thus it explains why we maintain that the results from the Poisson regression are not valid and concentrate on the results from the negative binomial regression in this study.

The results from several diagnostic tests have shown that model is well specified(see appendix C of study) Additionally, we can easily explain the results of the regression as semi-elasticities, hence measuring the relative variation of the conditional expected value for a variation of the ***i-th*** unit of the covariate, leaving other regressors constant. Table 1 shows that the APPP parameter for Denmark is negative and small in magnitude but statistically significant, indicating Swedish CPI has an effect on tourism demand from Denmark. The estimated absolute elasticity is – 0.70 % and greater than the other countries except Switzerland. This indicates that a 1% increase in *CPISW* results in a 0.7% increase in tourist arrivals to Sweden from Denmark. The low APPP elasticity for the US and UK could be a reflection of the appreciation of the Swedish krona against the US dollar and UK pound.

The estimated elasticity of the RPPP ranges from 1.6% to 8.8% and is greater than one for Japan and the US. This indicates that a 1% rise in the RPPP (price of tourism in Sweden relative to home country) causes a more than 1% fall in tourist arrivals from Japan and the US. These estimates indicate that tourist arrivals in Sweden from these countries are elastic with respect to the RPPP variable. This implies that Sweden must maintain its international price competitiveness to maintain high growth in tourist inflow. The estimated RPPP-level elasticity ranges from 0.2% to 0.8% and is more than one for Denmark and Japan. These figures suggest that a 1% increase in the RPPP results in a 0.2% and 0.8% decrease in tourist arrivals to the SW6 region from Denmark and Japan, respectively.

The high relative price elasticity for Japan may also be a reflection of the depreciation of the Swedish krona against the Japanese Yen. As expected, the estimated elasticities of RP_{tij} for the Denmark, UK, and the US are positive.

In the case of Denmark, we find that most dummies are significant, indicating clear seasonality in tourism demand. The demand in November is the highest for the year. In contrast, we find all the lags are not statistically significant. For the UK, the results showed most of the dummies are significant with the negative signs.

For Switzerland, only the summer dummies are large, positive and statistically significant, meaning that the Swiss are relatively more interested in summer tourism. The remaining dummies are either insignificant or small in magnitude. The estimated parameters of lags 1 and 11 are positive and significant.

In general, the lag of the dependent variable for the months of January is also significant, supporting the hypothesis of a habit-forming effect. Some of the monthly dummies as proxies for seasonal effects are also significant, including January, March, May, June, July, September, October and November. Estimates for Denmark and the US dummies show a clear seasonal variation in the pattern of Danish/American tourism demand for Sweden, such that demand in January, February, March and July, is higher than in October and November, with lower demand in other months. Additionally, estimates of the Japan dummy show demand is higher in February, March, October and November, with lower demand in other months.

5. Summary and Conclusions

The main purpose of this paper is to estimate the demand function for tourism to the SW6 region of Sweden from five different countries: Denmark, the UK, Switzerland, Japan and the US. Monthly time-series count data from 1993:01 to 2008:12 is collected from Statistics Sweden for this purpose. For each visiting country, we specify a separate equation with the relative information included in each equation. We estimate these equations using a TNBM, which takes into consideration the count-data-dependent variable.

The results show that APPP, RPPP, some lagged dependent variables and several monthly dummy variables representing seasonal effects, have a significant impact on the number of visitors to the SW6 region. The results also show that the RPPP and relative price exchange rate have a significant effect on international tourism demand from some countries. However, although we could view this conclusion as supporting a theoretical framework that describes the relationship of variables in the tourism demand model, our demand system lacks a travel-cost variable. Nonetheless, our results could also have important implications for the decision-making process of tourism agencies in Sweden when considering influential factors in their long-run planning.

Table 1: Negative Binomial Regression Estimation Results for Visitors to Sweden

Sweden	Equations				
Parameters	Denmark	UK	Switzerland	Japan	US
Constant	8.791978 (1.497309) P=0.0000	2.835711 (2.867543) P=0.3227	5.494186 (2.511995) P=0.0287	9.108329 (2.483148) P= 0.0002	6.425795 (3.310682) P=0.0523
APPP	-0.006964 (0.004269) P=0.0128	0.004116 (0.008375) P=0.6231	0.013039 (0.005238) P=0.0128	-0.004724 (0.005750) P=0.4114	0.003128 (0.007542) P=0.6783
RP_{tij}	0.014384 (0.007026) P=0.0406	0.174141 (0.030868) P=0.0004	-0.002774 (0.001462) P=0.0577	-0.010533 (0.050940) P=0.8362	0.029906 (0.057171) P=0.6009
RPPP	-0.168004 (0.156627) P=0.2834	0.389266 (0.319991) P=0.2238	-0.675348 (0.416469) P=0.1049	-0.883306 (0.367171) P=0.0161	-0.670766 (0.474972) P=0.1579
D1	1.456599 (0.356855) 4.081766 P =0.0000	0.319991 (0.163118) P=0.0164	-0.340746 (0.446777) P=0.4457	0.232495 (0.283324) P=0.4119	-0.130724 (0.232189) P=0.5734
D2	2.037094 (0.504795) P=0.0001	0.163118 (0.175903) P=0.1257	-0.593054 (0.614039) P=0.3341	0.166543 (0.270680) P=0.5384	-0.143697 (0.265009) P=0.5877
D3	1.099070 (0.554386) P=0.0474	0.175903 (0.188178) P=0.4962	-0.140470 (0.620564) P=0.8209	-0.033810 0.258547 0.8960	0.315219 (0.245348) P=0.1989

D4	-0.965434 (0.491610) P=0.0496	0.188178 (0.173503) P=0.0047	-0.941426 (0.546567) P=0.0850	-1.184628 (0.272251) P=0.0000	-0.447211 (0.227041) P=0.0489
D5	-2.103995 (0.440510) P=0.0000	0.173503 (0.177696) P=0.0000	-0.694128 (0.535532) P=0.1949	-1.176119 (0.302216) P= 0.0001	-0.509309 (0.228909) P=0.0261
D6	-0.229370 (0.460962) P=0.6188	0.177696 (0.167236) P=0.2082	0.852633 (0.556489) P=0.1255	-0.589441 (0.284675) P=0.0384	0.815213 (0.225461) P=0.0003
D7	1.357418 (0.450139) P=0.0026	0.167236 (0.177898) P=0.5845	1.841073 (0.536119) P=0.0006	-0.506495 (0.300574) P=0.0920	0.344866 (0.235538) P=0.1431
D8	0.056738 (0.474588) 0.9048	0.177898 (0.176815) P=0.0752	0.286989 (0.543042) P=0.5972	-0.292508 (0.270406) P=0.2794	-0.067937 (0.235435) P=0.7729
D9	-1.007417 (0.496794) 0.0426	0.176815 (0.190509) P= 0.0000	-1.622040 (0.612338) P=0.0081	-0.397019 (0.260878) P=0.1280	-0.476267 (0.245217) P=0.0521
D10	-1.296891 (0.476244) P=0.0065	0.190509 (0.181707) P=0.0000	-2.031443 (0.620912) P=0.0011	-1.029751 (0.270561) P=0.0001	-0.454839 (0.276272) P=0.0997
D11	-1.697296 (0.336548) P=0.0000	0.181707 (0.178382) P=0.0000	0.161492 (0.470259) P=0.7313	0.927412 (0.274075) P=0.0007	-0.121631 (0.231899) P=0.5999
$Y_{(t-1)}$	non significant	0.000187 (3.49E-05) P=0.0000	non significant	0.000403 (8.58E05) P=0.0000	0.000419 (8.32E-05) P=0.0000
$Y_{(t-2)}$	non significant	non significant	0.000397 (0.000174) P=0.0224	non significant	0.000229 (9.29E-05) P=0.0135
$Y_{(t-7)}$	non significant	non significant	non significant	0.000189 (7.93E-05) P=0.0173	non significant
$Y_{(t-11)}$	non significant	non significant	non significant	non significant	non significant
$Y_{(t-12)}$	non significant	non significant	non significant	0.000339 (8.69E-05) P=0.0001	non significant
R^2 Log likelihood Restr. log likelihood LR statistic (27 df) Probability(LR stat) LR index (Pseudo-R2) Meandependent var S.D. dependent var	0.955694 -1412.890 -1818855. 3634885 0.000000 0.999223 15937.71 21303.70	0.873208 -1161.288 -118091.1 233859.6 0.000000 0.990166 2008.449 1959.605	0.860487 -999.1994 -45228.28 88458.16 0.000000 0.977908 597.3462 681.6133	0.504328 -1052.794 -53192.29 104279.0 0.000000 0.980208 784.9487 812.2214	0.545994 -1100.147 -31210.32 -31210.32 0.000000 0.964751 851.6795 632.7129

Sources: Derived from tables 1 to 5 in the appendix of this study

References

Allison, P. D. and Waterman, R. (2002), 'Fixed effects negative binomial regression models', in Ross M. Stolzenberg (ed.), Sociological Methodology Oxford: Basil Blackwell.

Atkins, D.C. and Gallop, R.J. (2007), 'Rethinking how family researchers model infrequent outcomes:A tutorial on count regression and zero-inflated models',Journal of Family Psychology, vol. 21, pp. 726–735.

Green, W. H. (2003), Econometrics analysis, [where?]: PrenticeHall International.

Hausman, J., Hall, B. H. andGriliches, Z. (1984), 'Econometric models for count data with an application to the patents-R&D relationship', Econometrica, vol. 52, pp. 909–938.

Hultkrantz, L. (1995), 'Dynamic price response of inbound tourism guest nights in Sweden', Tourism Economics, vol. 1 (4), pp. 357–374.

Hultkrantz, L. and Olsson, C. (1997), 'Chernobyl effect on domestic and inbound tourism in Sweden', Environment and resources economics, vol.9, pp. 239–258.

Jorgensen and Solvoll G. (1996), 'Demand models for inclusive tour charter: The Norwegian Case', Tourism Management, vol. 17(1), pp. 17–24.

Kronenbery, K., (2013)Advertising effectiveness on international tourism demand I Åre- An Econometrics Analysis.Mid-Sweden University: European Tourism Research Centre (ETOUR), Department of Social Science, (unpublished) Master of Science

Long, J.S. and Freese , J. (2006) Regression models for categorical dependent variables using Stata, second edition, TX Stata Press.

Pilbeam, K. (2013) International Finance, UK: Palgrave Macmillan

Salman, A. K. (2003), Estimating tourist demand through co-integration analysis: Swedish data', Current Issues in Tourism, vol. 6(4). Pp332-339.

Salman, A. K. (2011), 'Using the SUR model of tourism demand for neighboring Regions in Sweden and Norway' in MiroslavVerbic (ed.) Advance in Econometrics; Theory and Applications, Croatia: INTECH Open Access Publisher,pp.97–116.ISBN 978-953-307- 503–7.

Salman, A. K., Shukur G. and Bergmann-Winberg, M. (2007), 'Comparison of econometric Modelling of demand for domestic and international tourism: Swedish data', Current Issues in Tourism, vol. 10(4). pp 332-342.

Salman, A. K., Sörensson, A., Arnesson, L. and Shukur, G. (2010), 'Estimating international tourism demand for selected regions in Sweden and Norway with iterative seemingly unrelated regression (ISUR)',Scandinavian Journal of Hospitality and Tourismvol. 10(4), pp. 395–410.

Sano, Y., Jeong, Y., Acock, A.C. and Zvonkovic,A.M. (2005) 'Working with count data: practical demonstration of Poisson, negative binomial, and zero-inflated regression models'. 35th Annual Theory Construction and Research Methodology Workshop – Proceedings, National Council of Family Relations,

Phoenix, AZ.Sturman, M.C. (1999), 'Multiple approaches to analyzing count data in studies of individual differences: The propensity for type I error, illustrated with the case of absenteeism prediction',Educational and Psychological Measurement, vol. 59, pp. 414–430.

Srivastava, V. and Giles, D. (1987) seemingly unrelated regression equations models, New York: Marcel Dekker.

Tabachnick, B.G. andFidell, L.S.(2007)Using multivariate statistics, 5th ed., Boston: Allyn and Bacon.

Verbeek, M. (2008) A guide to modern econometrics, London: John Wiley and Sons Ltd

Washington, S., M. Karlaftis, and F. Mannering (2003) Statistical and Econometics method for transportation Data analysis, Champman and Hall, Boca Raton.

Whitely, J.D. (1994) A course in macroeconomic modeling and forecasting, UK: Harvester Wheatsheaf.

Appendix A: Estimation results for the international demand function(Negative binomial regression model (TNBM)

Table 1. Negative Binomial Regression Model (TNBM) For Denmark

	Coefficient	Std. Error	z-Statistic	Prob.
Constant	8.791978	1.497309	5.871854	0.0000
Appp	-0.006964	0.004269	-1.631231	**0.0128**
RP_{tij}	0.014384	0.007026	2.047208	**0.0406**
Rppp	-0.168004	0.156627	-1.072635	0.2834
D1	1.456599	0.356855	4.081766	0.0000
D2	2.037094	0.504795	4.035484	0.0001
D3	1.099070	0.554386	1.982501	0.0474
D4	-0.965434	0.491610	-1.963821	0.0496
D5	-2.103995	0.440510	-4.776270	0.0000
D6	-0.229370	0.460962	-0.497590	0.6188
D7	1.357418	0.450139	3.015550	0.0026
D8	0.056738	0.474588	0.119552	0.9048
D9	-1.007417	0.496794	-2.027836	0.0426
D10	-1.296891	0.476244	-2.723163	0.0065
D11	-1.697296	0.336548	-5.043256	0.0000
Y1S(-1)	2.43E-06	5.97E-06	0.407124	0.6839
Y1S(-2)	1.17E-05	6.33E-06	1.842276	0.0654
Y1S(-3)	4.24E-06	6.22E-06	0.681202	0.4957
Y1S(-4)	-3.34E-06	6.23E-06	-0.536452	0.5916
Y1S(-5)	1.98E-06	5.59E-06	0.354087	0.7233
Y1S(-6)	9.98E-07	5.31E-06	0.187909	0.8509
Y1S(-7)	-1.55E-06	5.16E-06	-0.301177	0.7633
Y1S(-8)	-4.64E-06	5.37E-06	-0.863684	0.3878
Y1S(-9)	7.16E-06	5.34E-06	1.340974	0.1799
Y1S(-10)	-6.93E-09	5.65E-06	-0.001227	0.9990
Y1S(-11)	9.98E-06	5.58E-06	1.788024	0.0738
Y1S(-12)	5.54E-06	5.24E-06	1.056868	0.2906
	Mixture Parameter			
SHAPE:C(28)	-2.112733	0.111596	-18.93197	0.0000
R-squared	**0.955694**	Meandependent var		15937.71
Adjusted R-squared	**0.946348**	S.D. dependent var		21303.70
S.E. of regression	4934.537	Akaike info criterion		18.47295
Sumsquaredresid	3.12E+09	Schwarz criterion		19.02036
Log likelihood	-1412.890	Hannan-Quinn criter.		18.69528
Restr. log likelihood	-1818855.	Avg. log likelihood		-9.056986
LR statistic (27 df)	3634885.	LR index (Pseudo-R2)		0.999223
Probability(LR stat)	0.999223			

Table 2: Negative Binomial Regression Model (RENBM) For UK

	Coefficient	Std. Error	z-Statistic	Prob.
C	2.835711	2.867543	0.988899	0.3227
Appp	0.004116	0.008375	0.491473	0.6231
RP_{tij}	0.174141	0.030868	5.641493	**0.0000**
Rppp	-0.389266	0.319991	1.216491	0.2238
D1	-0.391616	0.163118	-2.400823	0.0164
D2	0.269372	0.175903	1.531364	0.1257
D3	-0.128056	0.188178	-0.680505	0.4962
D4	-0.489896	0.173503	-2.823566	0.0047
D5	-1.434248	0.177696	-8.071349	0.0000
D6	-0.210487	0.167236	-1.258622	0.2082
D7	-0.097281	0.177898	-0.546836	0.5845
D8	-0.314628	0.176815	-1.779417	0.0752
D9	-1.068363	0.190509	-5.607939	0.0000
D10	-1.421334	0.181707	-7.822112	0.0000
D11	-1.287411	0.178382	-7.217154	0.0000
Y2S(-1)	0.000187	3.49E-05	5.364053	0.0000
Y2S(-2)	1.27E-05	4.14E-05	0.306691	0.7591
Y2S(-3)	9.10E-07	4.21E-05	0.021623	0.9827
Y2S(-4)	1.52E-05	4.16E-05	0.365312	0.7149
Y2S(-5)	-1.10E-05	4.33E-05	-0.253843	0.7996
Y2S(-6)	1.99E-05	4.07E-05	0.488710	0.6250
Y2S(-7)	-4.30E-05	4.04E-05	-1.064703	0.2870
Y2S(-8)	3.23E-05	4.10E-05	0.789699	0.4297
Y2S(-9)	8.12E-05	4.47E-05	1.818130	0.0690
Y2S(-10)	3.93E-05	4.48E-05	0.877193	0.3804
Y2S(-11)	7.94E-05	4.49E-05	1.769047	0.0769
Y2S(-12)	-4.89E-05	3.73E-05	-1.312202	0.1895
	Mixture Parameter			
SHAPE:C(28)	-2.416885	0.113025	-21.38364	0.0000

R-squared	0.873208	Meandependent var	2008.449
Adjusted R-squared	0.846462	S.D. dependent var	1959.605
S.E. of regression	767.8495	Akaike info criterion	15.24728
Sumsquaredresid	75467888	Schwarz criterion	15.79469
Log likelihood	-1161.288	Hannan-Quinn criter.	15.46962
Restr. log likelihood	-118091.1	Avg. log likelihood	-7.444154
LR statistic (27 df)	233859.6	LR index (Pseudo-R2)	0.990166
Probability(LR stat)	0.000000		

Table 3. Negative Binomial Regression Model (RENBM) for Swaziland

	Coefficient	Std. Error	z-Statistic	Prob.
C	5.494186	2.511995	2.187180	0.0287
Appp	0.013039	0.005238	2.489165	**0.0128**
RP_{tij}	-0.002774	0.001462	-1.897878	**0.0577**
Rppp	-0.675348	0.416469	-1.621603	0.1049
D1	-0.340746	0.446777	-0.762676	0.4457
D2	-0.593054	0.614039	-0.965824	0.3341
D3	-0.140470	0.620564	-0.226358	0.8209
D4	-0.941426	0.546567	-1.722434	0.0850
D5	-0.694128	0.535532	-1.296146	0.1949
D6	0.852633	0.556489	1.532165	0.1255
D7	1.841073	0.536119	3.434075	0.0006
D8	0.286989	0.543042	0.528484	0.5972
D9	-1.622040	0.612338	-2.648929	0.0081
D10	-2.031443	0.620912	-3.271711	0.0011
D11	0.161492	0.470259	0.343411	0.7313
Y3S(-1)	0.000253	0.000167	1.513239	0.1302
Y3S(-2)	0.000397	0.000174	2.283142	0.0224
Y3S(-3)	0.000101	0.000175	0.575637	0.5649
Y3S(-4)	2.89E-05	0.000179	0.161947	0.8713
Y3S(-5)	-0.000134	0.000175	-0.768977	0.4419
Y3S(-6)	-0.000179	0.000169	-1.060486	0.2889
Y3S(-7)	0.000308	0.000172	1.787535	0.0739
Y3S(-8)	-4.57E-05	0.000168	-0.272834	0.7850
Y3S(-9)	0.000161	0.000161	0.995560	0.3195
Y3S(-10)	-0.000123	0.000159	-0.775161	0.4382
Y3S(-11)	-2.00E-05	0.000169	-0.118227	0.9059
Y3S(-12)	-2.96E-05	0.000161	-0.184317	0.8538
	Mixture Parameter			
SHAPE:C(28)	-1.874802	0.114742	-16.33921	0.0000

R-squared	0.860487	Meandependent var	597.3462
Adjusted R-squared	0.831059	S.D. dependent var	681.6133
S.E. of regression	280.1597	Akaike info criterion	13.16922
Sumsquaredresid	10046651	Schwarz criterion	13.71663
Log likelihood	-999.1994	Hannan-Quinn criter.	13.39156
Restr. log likelihood	-45228.28	Avg. log likelihood	-6.405125
LR statistic (27 df)	88458.16	LR index (Pseudo-R2)	0.977908
Probability(LR stat)	0.000000		

Table4. Negative Binomial Regression Model (RENBM) For Japan

	Coefficient	Std. Error	z-Statistic	Prob.
C	9.108329	2.483148	3.668057	0.0002
Appp	-0.004724	0.005750	-0.821430	0.4114
RP_{tij}	-0.010533	0.050940	-0.206769	0.8362
Rppp	-0.883306	0.367171	-2.405708	0.0161
D1	0.232495	0.283324	0.820596	0.4119
D2	0.166543	0.270680	0.615277	0.5384
D3	-0.033810	0.258547	-0.130769	0.8960
D4	-1.184628	0.272251	-4.351232	0.0000
D5	-1.176119	0.302216	-3.891649	0.0001
D6	-0.589441	0.284675	-2.070578	0.0384
D7	-0.506495	0.300574	-1.685096	0.0920
D8	-0.292508	0.270406	-1.081736	0.2794
D9	-0.397019	0.260878	-1.521854	0.1280
D10	-1.029751	0.270561	-3.805984	0.0001
D11	0.927412	0.274075	3.383792	0.0007
Y4S(-1)	0.000403	8.58E-05	4.693634	0.0000
Y4S(-2)	-2.29E-05	8.43E-05	-0.271319	0.7861
Y4S(-3)	0.000136	8.38E-05	1.622586	0.1047
Y4S(-4)	4.63E-05	8.04E-05	0.576254	0.5644
Y4S(-5)	-1.46E-05	7.68E-05	-0.190703	0.8488
Y4S(-6)	-4.11E-05	7.74E-05	-0.530754	0.5956
Y4S(-7)	0.000189	7.93E-05	2.381208	0.0173
Y4S(-8)	9.91E-05	8.45E-05	1.171999	0.2412
Y4S(-9)	-8.05E-05	7.86E-05	-1.023835	0.3059
Y4S(-10)	2.81E-05	7.45E-05	0.377481	0.7058
Y4S(-11)	2.79E-06	8.76E-05	0.031882	0.9746
Y4S(-12)	0.000339	8.69E-05	3.901100	0.0001
	Mixture Parameter			
SHAPE:C(28)	-1.779961	0.112142	-15.87241	0.0000
R-squared	0.504328	Meandependent var		784.9487
Adjusted R-squared	0.399772	S.D. dependent var		812.2214
S.E. of regression	629.2636	Akaike info criterion		13.85634
Sumsquaredresid	50684497	Schwarz criterion		14.40375
Log likelihood	-1052.794	Hannan-Quinn criter.		14.07867
Restr. log likelihood	-53192.29	Avg. log likelihood		-6.748681
LR statistic (27 df)	104279.0	LR index (Pseudo-R2)		0.980208
Probability(LR stat)	0.000000			

Table5: Negative Binomial Regression Model (RENBM) For United States

	Coefficient	Std. Error	z-Statistic	Prob.
C	6.425795	3.310682	1.940928	0.0523
Appp	0.003128	0.007542	0.414731	0.6783
RP_{tij}	-0.029906	0.057171	0.523090	0.6009
Rppp	-0.670766	0.474972	-1.412223	0.1579
D1	-0.130724	0.232189	-0.563009	0.5734
D2	-0.143697	0.265009	-0.542233	0.5877
D3	0.315219	0.245348	1.284781	0.1989
D4	-0.447211	0.227041	-1.969738	0.0489
D5	-0.509309	0.228909	-2.224940	0.0261
D6	0.815213	0.225461	3.615767	0.0003
D7	0.344866	0.235538	1.464163	0.1431
D8	-0.067937	0.235435	-0.288558	0.7729
D9	-0.476267	0.245217	-1.942226	0.0521
D10	-0.454839	0.276272	-1.646346	0.0997
D11	-0.121631	0.231899	-0.524501	0.5999
Y5S(-1)	0.000419	8.32E-05	5.038315	0.0000
Y5S(-2)	0.000229	9.29E-05	2.469782	0.0135
Y5S(-3)	4.10E-05	0.000113	0.361624	0.7176
Y5S(-4)	-6.07E-05	0.000102	-0.595790	0.5513
Y5S(-5)	-6.46E-06	0.000105	-0.061360	0.9511
Y5S(-6)	0.000142	0.000105	1.353065	0.1760
Y5S(-7)	3.83E-05	0.000105	0.366586	0.7139
Y5S(-8)	-0.000148	0.000102	-1.448242	0.1475
Y5S(-9)	0.000100	0.000101	0.985771	0.3242
Y5S(-10)	9.67E-05	9.97E-05	0.970112	0.3320
Y5S(-11)	5.47E-05	9.31E-05	0.587379	0.5569
Y5S(-12)	-6.69E-05	8.53E-05	-0.784321	0.4329
	Mixture Parameter			
SHAPE:C(28)	-1.801258	0.111594	-16.14113	0.0000
R-squared	0.545994	Meandependent var		851.6795
Adjusted R-squared	0.450227	S.D. dependent var		632.7129
S.E. of regression	469.1356	Akaike info criterion		14.46343
Sumsquaredresid	28171288	Schwarz criterion		15.01084
Log likelihood	-1100.147	Hannan-Quinn criter.		14.68576
Restr. log likelihood	-31210.32	Avg. log likelihood		-7.052227
LR statistic (27 df)	60220.34	LR index (Pseudo-R2)		0.964751
Probability(LR stat)	0.000000			

Appendix B

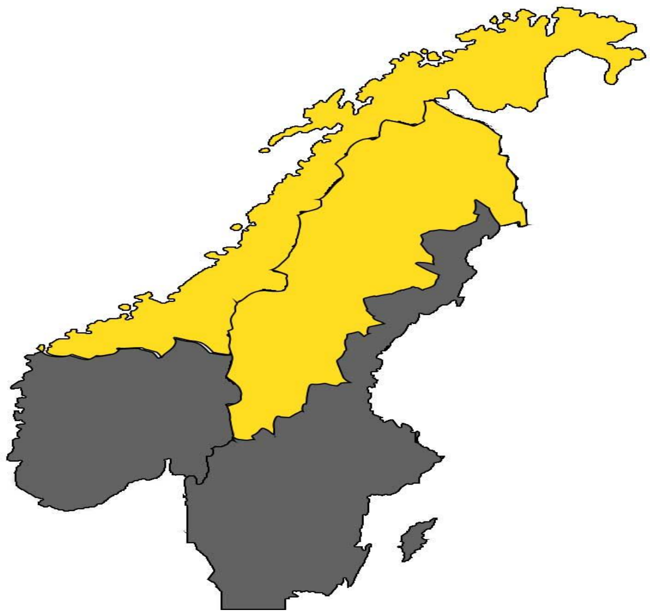


Figure 1: Swedish and Norwegian Maps; the Objective 6 region (SW: 6) in Sweden is the lightly shadowed area at the top and top-left of the map of Sweden. The North –Norway included and Tröndelag region in Norway (NWT) is the lightly shadowed part on the top right of the map of Norway

Appendix C

Diagnostic Tests

The Cusum test

This test is used for time series and checks for structural changes. In the Cusum test Recursive Residuals (RR) calculated by the Kalman Filter are used.

I now describe the construction of recursive residuals and the Kalman filter technique. The recursive residuals can be computed by forward or backward recursion. Only forward recursion is described, backward recursion being analogous.

Given N observations, consider the linear model (2 . 2 . 1) but with the corresponding vector of coefficient □ expressed as $□_t$, implying that the coefficients may vary over time t. The hypothesis to be tested is □□□= □□□□= , . . ., =□□□□□□= □. The OLS estimator based on N observations is: $b = (X'X)^{-1} X'y$,

where X is a N by k matrix of observations on the regressors, and y is an N by 1 vector of observations for the dependent variable. Suppose that only r observations are used to estimate □. Then for r > k, where k is the number of independent variables,

$$b_r = (X_r'X_r)^{-1} X_r'y_r, \; r = k+1, \ldots, N .$$

Using b_r, one may "forecast" y_r at sample point r, corresponding to the vector X_r of the explanatory variables at that point.

Recursive residuals are now derived by estimating equation (2 . 2 . 1) recursively in the same manner, that is by using the first k observations to get an initial estimate of □, and then gradually enlarging the sample, adding one observation at a time and re-estimating □ at each step. In this way, it is possible to get (N-k) estimates of the vector □, and correspondingly (N-k-1) forecast errors of the type: $W_r = y_r - X_r b_{r-1}$, r = k+1, . . ., N , where b_{r-1} is an estimate of based on the first r - 1 observations. It can be shown that, under the null hypothesis, these forecast errors have mean zero and variance^2 d_r^2, where d_r is a scalar function of the explanatory variables, equal to $[1 + X_r'(X'_{r-1}X_{r-1})^{-1} X_r]^{1/2}$.

Then the quantity: $$W_r = \frac{y_r - X_r b_{r-1}}{\left[1+ X_r'(X_{r-1}'X_{r-1})X_r \right]^{1/2}}, \quad r = k+1, \ldots, N,$$

gives a set of standardized prediction errors, called "recursive residuals". The recursive residuals are independently and normally distributed with mean zero and constant variance ^{2}. As a result of a change in the structure over time, these recursive residuals will no longer have zero mean, and the CUSUM of these residuals can be used to test for structural change.

CUSUM involves the plot of the quantity: $$V_r = \sum_{t=k+1}^{r} W_t / \sigma^*, \quad r = k+1, \ldots, N,$$

where * is the estimated standard deviation based on the full sample.

The test finds parameter instability if the cumulative sum goes outside the area between the two error bounds. Thus, movements of V_t outside the error bounds are a sign of parameter instability.

The Breusch-Godfrey-test

The Breusch-Godfrey test can be separated into several stages:

1. Run an OLS on:

$$y_t = \alpha + \beta X_t + \theta_i y_{t-i} + \varepsilon_t$$

This gives us $\hat{\varepsilon}_t$

2. Run an OLS on:

$$\hat{\varepsilon}_t = \alpha + \beta X_t + \theta_i y_{t-i} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \ldots + \rho_P \hat{\varepsilon}_{t-P} + u_t$$

This equation can be used for any AR(P) process. From this equation the unrestricted residual sum of squares (RSS_u).

The restricted residual sum of squares (RSS_R) is given from the following equation:

$$\hat{\varepsilon}_t = \alpha + \beta X_t + \theta y_{t-1} + \nu_t$$

The null hypothesis is:

$$H_0 : \rho_1 = \rho_2 = \ldots. = \rho_P = 0$$

3. Run an F-test:

$F=((RSS_R-RSS_U)/p) / (RSS_U/(T-k-P))$

This has a distribution: F(P,T-k-P) under the null hypothesis.

The Breusch-Godfrey test can be tested for AR(P) processes which gives this test a clear advantage over other available tests for autocorrelation.

The Ramsey RESET-test

RESET test stands for Regression Specification Error Test. The test is very general and can only tell you if you have a problem or not. It tests for omitted variables and incorrect functional forms or misspecified dynamics and also if there is a correlation between the error term and the independent variable. The null hypothesis is:

H_0: E (ε_i/X_i) = 0
H_1: E (ε_i/X_i) ≠ 0 (and an omitted variable effect is present)

Thus, by rejecting the null hypothesis indicates some type of misspecification. First a linear regression is specified:

$$y_i = \alpha + \beta X_i + \varepsilon_i$$

This gives the restricted residual sum of squares (RSSR). After the RSS_R has been found the unrestricted model is presented by adding variables (three fitted values):

$$y_t = \alpha + \beta X_i + \theta_1 \hat{y}_i^2 + \theta_2 \hat{y}_i^3 + u_t$$

This gives us the unrestricted residual sum of squares (RSS_U). In the third step the RESET-test uses a F-test:
F= ((RSS_R - RSS_U)/number of restrictions under H_0) / (RSS_U / (N- number of parameters in unrestricted model))

The F-test checks if θ_1=θ_2 =0, if θ_1=θ_2≠0 I have an omitted variable or a misspecification in the model.
The White's test

This test is a general test where I do not need to make any specific assumptions regarding the nature of the heteroscedasticity, whether it is increasing, decreasing etc. The test only tells us if I have an indication of heteroscedasticity.

$$H_0 : \sigma_i^2 = \sigma^2 \qquad \forall i$$

The alternative hypothesis is not H_0, anything other than H_0.

The test can be divided into several steps:

1. Run an OLS on: $y_i = \alpha + \beta_1 X_{1i} + ... + \beta_k X_{ki} + \varepsilon_i$

From this equation I get $\hat{\varepsilon}_i$ which is used as a proxy for the variance.

2. Run an OLS on: $\hat{\varepsilon}_i^2 = \alpha_0 + \alpha_1 X_{1i} + ... + \alpha_k X_{ki} + \alpha_{k+1} X_{1i}^2 + ... + \alpha_{k+k} X_{k+k}^2 + \alpha_{k+k+1} X_{1i} X_k + \delta_i$
   Where k is the number of parameters. The variance is considered to be a linear function of a number of independent variables, their quadratic and cross products. Thus, the X:s is used as a proxy for Z.
3. Calculate an F-test:

Restricted model: $\hat{\varepsilon}_i^2 = \alpha_0^{'} + \delta_i^{'}$

From this test the restricted residual sum of squares (RSS_R) is measured.

The F-test is:

F=((RSS_R-RSS_U)/k) / (RSS_U/(n-k-1))
Where $H_0 : \alpha_i = 0 \ \forall i = 1,2...k$

The ARCH Engel's LM test

This is a test for AutoRegressive Conditional Heteroscedasticity (ARCH). The ARCH process can be modeled as:
$y_t = \alpha + \beta X_t + \varepsilon_t$

where the Variance of ε_t conditioned on ε_{t-i} : Var($\varepsilon_t \backslash \varepsilon_{t-i}$) = $\alpha_0 + \alpha_1 \varepsilon_{t-i}^2$

1) Use OLS on the original model and get: $\hat{\varepsilon}_t$. Square it and use it in the following unrestricted model:

2) $$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_i \hat{\varepsilon}_{t-i}^2 + \delta_t$$

3) Test whether α_i = 0, for any i = 1, 2 , . . . By an F-test as before.

Test for Non-Normality

The test for non-normality is normally done before one test for heteroskedasticity and structural changes.

The test used here for testing for normal distribution is the Jarque-Bera test. The Jarque-Bera test is structured as follows:

$$T\left[1/6\hat{b}_1^2 + 1/24(\hat{b}_2 - 3)^2\right]$$
$$b_1 = \mu_3 /(\mu_2)^{3/2}$$
$$b_2 = \mu_4 /(\mu_2)^2$$

Where T is the total number of observations, b_1 is a measure for skewness and b_2 is a measure for kurtosis. The μ are different moments. The test has a chi-square distribution with two degrees of freedom under the null hypothesis of normal distribution. The two degrees of freedom comes from having one for skewness and one for kurtosis.

Single Equation Estimation and Diagnostic Results

Equation 1 (Denmark)

Equation 1

Dependent Variable: LY1S

Method: Least Squares

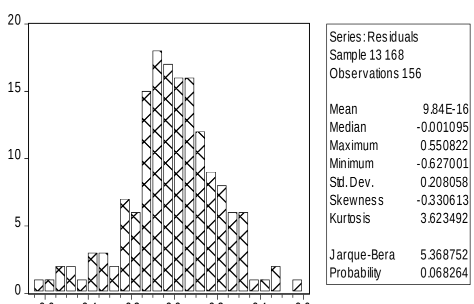
Date: 12/22/07 Time: 01:02

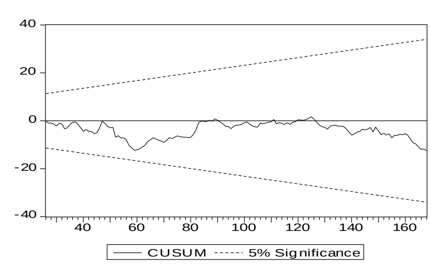
Sample(adjusted): 13 168

Included observations: 156 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.119869	2.177339	-0.055053	0.9562
LX11S	-0.329812	0.859238	-0.383842	0.7017
LX21S	1.949515	0.879789	2.215890	0.0283
LX31S	-0.174692	0.179727	-0.971984	0.3327
D1	0.535168	0.113648	4.708980	0.0000
D2	0.575040	0.164450	3.496738	0.0006
D3	0.193673	0.169505	1.142578	0.2552
D4	-0.441525	0.143971	-3.066755	0.0026
D5	-1.009611	0.125585	-8.039258	0.0000
D6	-0.232609	0.069912	-3.327157	0.0011
D7	0.258146	0.093145	2.771445	0.0063
D8	-0.408344	0.133995	-3.047455	0.0028
D9	-0.784962	0.111768	-7.023144	0.0000
D10	-0.803996	0.097960	-8.207415	0.0000
D11	-0.865023	0.102241	-8.460669	0.0000
LY1S(-1)	0.156265	0.083752	1.865806	0.0642
LY1S(-12)	0.092566	0.079332	1.166823	0.2453

R-squared	0.941591	Meandependent var	3.766018
Adjusted R-squared	0.934867	S.D. dependent var	0.666149
S.E. of regression	0.170009	Akaike info criterion	-0.603369
Sumsquaredresid	4.017505	Schwarz criterion	-0.271013
Log likelihood	64.06280	F-statistic	140.0472
Durbin-Watson stat	2.026049	Prob(F-statistic)	0.000000





Breusch-Godfrey Serial Correlation LM Test: (lag 1)

F-statistic	2.453516	Probability	0.119488

Breusch-Godfrey Serial Correlation LM Test: (lag 12)

F-statistic	1.187923	Probability	0.298218

White Heteroskedasticity Test:

F-statistic	1.560470	Probability	0.087545

Ramsey RESET Test:

F-statistic	2.804172	Probability	0.063938

ARCH Test: (1 lag)

F-statistic	0.669005	Probability	0.414671

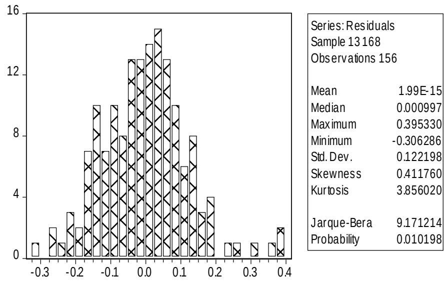
ARCH Test: (12 lag)

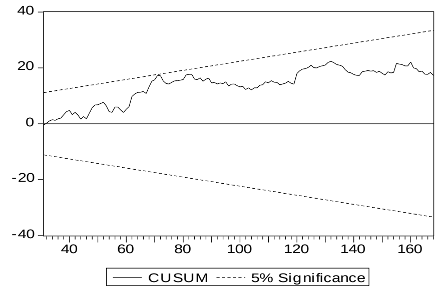
F-statistic	1.277820	Probability	0.239016

Equation 2. (UK)
DependentVariable: LY2S
Method: Least Squares
Date: 12/22/07 Time: 01:15
Sample(adjusted): 13 168
Included observations: 156 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-7.638103	3.820955	-1.999004	0.0476
LX12S	3.274332	1.484819	2.205207	0.0291
LX22S	0.885046	0.430066	2.057930	0.0415
LX32S	0.557153	0.668809	0.833053	0.4063
D1	-0.356825	0.065210	-5.471938	0.0000
D2	-0.250278	0.066938	-3.738970	0.0003
D3	-0.395741	0.071868	-5.506512	0.0000
D4	-0.433470	0.065740	-6.593749	0.0000
D5	-0.722577	0.072784	-9.927690	0.0000
D6	-0.101814	0.057072	-1.783959	0.0766
D7	-0.311841	0.058173	-5.360573	0.0000
D8	-0.402540	0.058301	-6.904539	0.0000
D9	-0.676417	0.072282	-9.358021	0.0000
D10	-0.535941	0.065281	-8.209807	0.0000
D11	-0.348392	0.060646	-5.744678	0.0000
LY2S(-1)	0.579430	0.064477	8.986605	0.0000
LY2S(-4)	-0.143180	0.058915	-2.430272	0.0164
LY2S(-12)	0.214971	0.063376	3.391974	0.0009

R-squared	0.901984	Meandependent var	3.132188
Adjusted R-squared	0.889909	S.D. dependent var	0.390316
S.E. of regression	0.129507	Akaike info criterion	-1.142004
Sumsquaredresid	2.314527	Schwarz criterion	-0.790098
Log likelihood	107.0763	F-statistic	74.70182
Durbin-Watson stat	1.976678	Prob(F-statistic)	0.000000





Breusch-Godfrey Serial Correlation LM Test: (1 lag)

F-statistic	0.000310	Probability	0.985967

Breusch-Godfrey Serial Correlation LM Test: (lag 12)

F-statistic	0.454242	Probability	0.937245

White Heteroskedasticity Test:

F-statistic	1.611124	Probability	0.050065

Ramsey RESET Test:

F-statistic	0.506977	Probability	0.603447

ARCH Test: (1 lag)

F-statistic	0.344502	Probability	0.558107

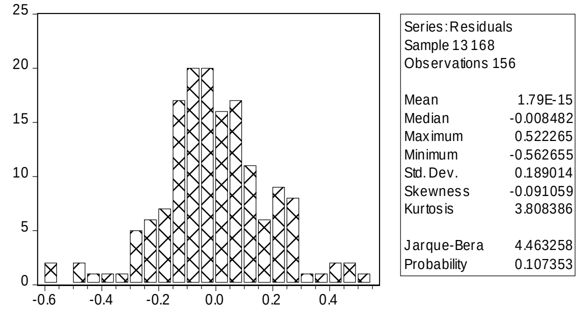
ARCH Test: (12)

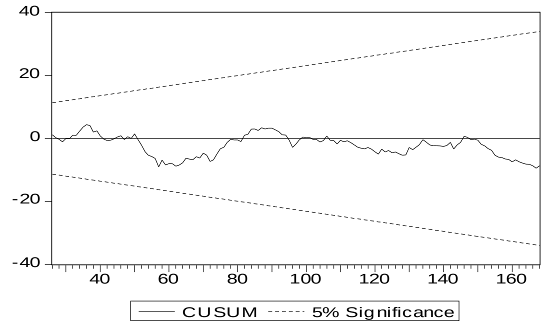
F-statistic	0.372600	Probability	0.971022

Equation 3 (Switzerland)
DependentVariable: LY3S
Method: Least Squares
Date: 12/26/07 Time: 13:20
Sample(adjusted): 13 168
Included observations: 156 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.161667	5.066274	0.229294	0.8190
LX13S	2.195696	1.541189	1.424677	0.1564
LX23S	-1.455462	0.946573	-1.537613	0.1264
LX33S	-2.230775	1.000473	-2.229722	0.0273
D3	0.214337	0.064273	3.334782	0.0011
D5	-0.108238	0.062189	-1.740476	0.0839
D6	0.463586	0.075057	6.176465	0.0000
D7	0.778064	0.089632	8.680650	0.0000
D8	0.445411	0.085466	5.211553	0.0000
D10	-0.342893	0.066463	-5.159183	0.0000
D11	0.300770	0.073663	4.083063	0.0001
LY3S(-1)	0.108503	0.059123	1.835206	0.0686
LY3S(-12)	0.186613	0.061763	3.021441	0.0030

R-squared	0.825277	Meandependent var	2.552475
Adjusted R-squared	0.810614	S.D. dependent var	0.452187
S.E. of regression	0.196785	Akaike info criterion	-0.333759
Sumsquaredresid	5.537558	Schwarz criterion	-0.079604
Log likelihood	39.03318	F-statistic	56.28637
Durbin-Watson stat	1.835514	Prob(F-statistic)	0.000000





Breusch-Godfrey Serial Correlation LM Test: (1 lag)

F-statistic	1.736072	Probability	0.189759

Breusch-Godfrey Serial Correlation LM Test: (lag 12)

F-statistic	0.631462	Probability	0.812301

White Heteroskedasticity Test:

F-statistic	1.254121	Probability	0.231790

Ramsey RESET Test:

F-statistic	4.581138	Probability	0.001663

ARCH Test: (1 lag)

F-statistic	4.51E-05	Probability	0.994649

ARCH Test: (12)

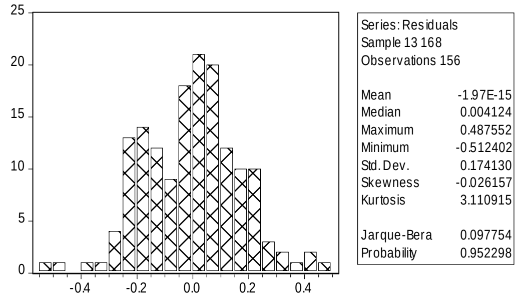
F-statistic	1.137238	Probability	0.335894

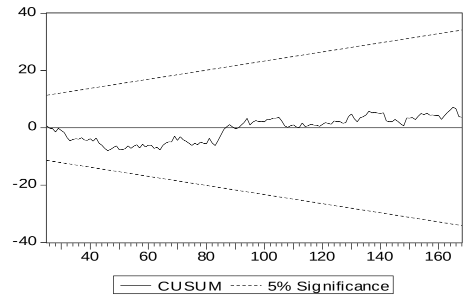
Equation 4 (Japan)

DependentVariable: LY4S
Method: Least Squares
Date: 12/26/07 Time: 13:46
Sample(adjusted): 13 168
Included observations: 156 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	16.92918	4.042619	4.187677	0.0000
LX14S	-6.459596	1.508936	-4.280895	0.0000
LX24S	0.097664	0.338176	0.288796	0.7732
LX34S	-2.374048	0.815590	-2.910834	0.0042
D2	0.158504	0.057138	2.774067	0.0063
D3	0.081410	0.057596	1.413475	0.1597
D4	-0.203459	0.063137	-3.222509	0.0016
D5	-0.140453	0.059189	-2.372963	0.0190
D6	0.119331	0.064106	1.861454	0.0647
D11	0.291585	0.076227	3.825223	0.0002
LY4S(-1)	0.233591	0.058468	3.995189	0.0001
LY4S(-12)	0.580193	0.053634	10.81761	0.0000

R-squared	0.845582	Meandependent var	2.684916
Adjusted R-squared	0.833786	S.D. dependent var	0.443123
S.E. of regression	0.180659	Akaike info criterion	-0.510612
Sumsquaredresid	4.699804	Schwarz criterion	-0.276008
Log likelihood	51.82774	F-statistic	71.68462
Durbin-Watson stat	2.090491	Prob(F-statistic)	0.000000





Breusch-Godfrey Serial Correlation LM Test: (1 lag)

F-statistic	0.529381	Probability	0.468057

Breusch-Godfrey Serial Correlation LM Test: (lag 12)

F-statistic	0.954636	Probability	0.495384

White Heteroskedasticity Test:

F-statistic	1.308870	Probability	0.199895

Ramsey RESET Test:

F-statistic	0.453596	Probability	0.636257

ARCH Test: (1 lag)

F-statistic	0.001458	Probability	0.969592

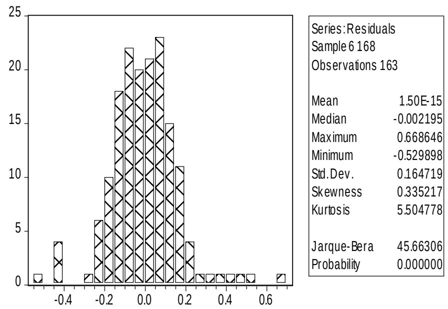
ARCH Test: (12)

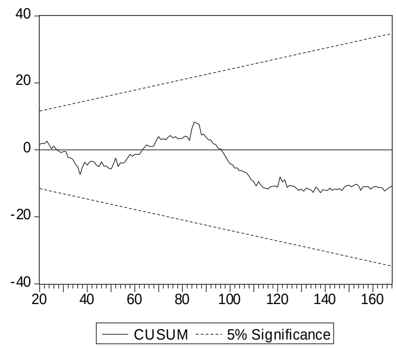
F-statistic	0.903619	Probability	0.545306

Equation 5 (USA)
DependentVariable: LY5S
Method: Least Squares
Date: 12/27/07 Time: 01:08
Sample(adjusted): 6 168
Included observations: 163 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6.162601	4.368918	-1.410555	0.1605
LX15S	2.690216	1.642524	1.637855	0.1036
LX25S	0.539904	0.395854	1.363897	0.1747
LX35S	0.217864	0.895073	0.243404	0.8080
D3	0.103305	0.057829	1.786378	0.0761
D4	-0.282404	0.057242	-4.933514	0.0000
D5	-0.136116	0.057171	-2.380880	0.0185
D6	0.335615	0.062417	5.376959	0.0000
D9	-0.305832	0.066889	-4.572240	0.0000
D10	-0.302530	0.068226	-4.434242	0.0000
D11	-0.112570	0.058869	-1.912222	0.0578
LY5S(-1)	0.599894	0.056822	10.55751	0.0000
LY5S(-3)	0.199785	0.074313	2.688433	0.0080
LY5S(-5)	-0.106157	0.066952	-1.585568	0.1150

R-squared	0.759922	Meandependent var	2.804902
Adjusted R-squared	0.738975	S.D. dependent var	0.336176
S.E. of regression	0.171754	Akaike info criterion	-0.603529
Sumsquaredresid	4.395427	Schwarz criterion	-0.337808
Log likelihood	63.18764	F-statistic	36.27932
Durbin-Watson stat	2.151594	Prob(F-statistic)	0.000000





Breusch-Godfrey Serial Correlation LM Test: (1 lag)

F-statistic	1.689666	Probability	0.195666

Breusch-Godfrey Serial Correlation LM Test: (lag 12)

F-statistic	0.610817	Probability	0.830248

White Heteroskedasticity Test:

F-statistic	1.203814	Probability	0.262503

Ramsey RESET Test:

F-statistic	2.500256	Probability	0.085549

ARCH Test: (1 lag)

F-statistic	6.588216	Probability	0.011182

ARCH Test: (12)

F-statistic	1.455332	Probability	0.148504