Museum and monument attendance and tourism flow: A time series analysis approach.

Cellini, Roberto and Cuccia, Tiziana

Dipartimento di economia e metodi quantitativi - Università di Catania

November 2009

MUSEUM & MONUMENT ATTENDANCE AND TOURISM FLOW:
A Time Series Analysis Approach

Roberto Cellini
University of Catania, Italy
Tiziana Cuccia
University of Catania, Italy

Abstract - This paper takes a time series analysis approach to evaluate the directions of causality between tourism flows, on the one side, and museum and monument attendance, on the other. We consider Italy as a case study, and analyze monthly data over the period January 1996 to December 2007. All considered series are seasonally integrated, and co-integration links emerge. We focus on the error correction mechanism among co-integrated time series to detect the directional link(s) of causality. Clear-cut results emerge: generally, the causality runs from tourist flows to museum and monument attendance. The non-stationary nature of time series, their co-integration relationships, and the direction of causal links suggest specific implication for tourism and cultural policies.

Keywords: Tourism, Museum, Seasonal unit root, Co-integration, Causality.

Roberto Cellini - Faculty of Economics, University of Catania (corso Italia, 55 – 95128 Catania, Italy. Email <cellini@unict.it>).
Tiziana Cuccia - Faculty of Economics, University of Catania. (corso Italia, 55 – 95128 Catania, Italy. Email <cucciati@unict.it>).
.

MUSEUM & MONUMENT ATTENDANCE AND TOURISM FLOW:
A Time Series Analysis Approach

Abstract: This paper takes a time series analysis approach to evaluate the directions of causality between tourism flows, on the one side, and museum and monument attendance, on the other. We consider Italy as a case study, and analyze monthly data over the period January 1996 to December 2007. All considered series are seasonally integrated, and co-integration links emerge. We focus on the error correction mechanism among co-integrated time series to detect the directional link(s) of causality. Clear-cut results emerge: generally, the causality runs from tourist flows to museum and monument attendance. The non-stationary nature of time series, their co-integration relationships, and the direction of causal links suggest specific implication for tourism and cultural policies.
Keywords: Tourism, Museum, Seasonal unit root, Co-integration, Causality.

1. INTRODUCTION

The effectiveness of cultural attractions in enhancing tourism flows is a widely debated issue, both in cultural economics and in tourism economics. Quite surprisingly, however, the evidence about the relationships between attendance at cultural attractions and tourist flows is restricted to specific, albeit interesting, cases, and general analyses are lacking, to the best of our knowledge. In particular, we did not find any single study analyzing the relationship at national levels, nor employing aggregate data over long periods of time. In this paper we show that it is possible to derive clear-cut results from the time series analysis of monthly time series concerning tourism flows and museum and monument attendance, taking Italy as a case study. More specifically, we are interested in studying which is the possible direction of causal links between tourism flows and the attendance at museums and monuments.

Not surprisingly, time series in the field of both tourism and cultural sites attendance show a great deal of seasonality; more specifically, all the time series taken into consideration in the present paper appear to be seasonally integrated. Hence, the techniques related to integration and co-integration among time series provide a natural language to study the relationships and, more specifically, the causal links among tourism flows and cultural sites attendance. It is worth stressing that, while the presence of seasonal unit roots in time series related to tourism is not a novelty in the literature, the present paper represents –as far as we know– the first attempt to investigate the causal direction between tourism flows and monument attendance, using time series analysis techniques based on seasonal co-integration.

The knowledge of the properties of time series, and their relationships, can shed light on the sound evaluation of different points: in the case at hand, we can derive implications on the effectiveness of cultural sites to attract tourism, on the effects of tourism dynamics upon the attendance at museums and monuments, and on the role of monuments and museums to lessen seasonality in tourism flows, just to mention the most prominent and obvious ones.

We anticipate that tourist presence as measured by overnight stays (as well as tourist arrivals and average stay) co-integrate with the attendance at museums and monuments, and a unidirectional long-run causal link generally does emerge, running from tourism flow variables to cultural sites attendance. Technically, tourism flows Granger-cause cultural sites attendance, while the reverse does not hold. Appropriate elasticity coefficients are estimated at the end. Consistently, it is hard to sustain that cultural attractions can promote tourism in the long run, at least in the aggregate. On the other hand, the long-run dynamics of visits to cultural sites is strongly determined by the dynamics of tourism flows. Therefore, we can guess that the role of cultural sites is limited in lessening seasonality.

There is a wide body of studies, especially in cultural economics, concerning museums. The largest part of this literature focuses on the microeconomic determinants of museum visits, on the demand side, and on problems of organization and governance on the supply side. An update review is provided by Frey and Meier (2006), in their chapter in the *Handbook of Cultural Economics* devoted to the economics of museums. Interestingly enough, however, their review on applied works cannot list any study, in which tourism flow is considered among the determinants of attendance at museums. In this perspective, our present exercise can complement the evidence in existing literature. Of course, we are aware that museums and monuments have values that go far beyond tourism motivation. Nevertheless, omitting tourism flows from the determinants of cultural visits can lead to serious mistakes, especially in relation to (here documented) cases of cointegrating relationships. Similarly, we are also aware that the visits to monuments and museums represent too strict a measure of cultural tourism, and hence our evidence –that causal links go from tourism flows to museum visits and not in the reverse sense– cannot justify the conclusion that "cultural tourism" is not important for enhancing tourism flows.

2. TOURISM FLOWS AND ATTENDANCE AT MUSEUMS AND MONUMENTS

Tourism studies seem to take for granted that museums and monuments play a significant role as tourist attractions. The available economic literature on museums, in particular, can be divided into two lines of research. The first line looks at the museum as an institution that has a private and a social role: it has to satisfy the present visitors but it has to improve and preserve the present collection for future generations (Johnson, 2003). The main issues within this research line concern the organizational form (e.g., Fedeli and Santoni, 2006), the managerial aspects of supplying services and merchandise, along with the strategies for reducing production costs and fund raising (Kotler and Kotler, 1998). A large number of empirical studies are devoted to estimating visitors' willingness to pay and their price elasticity, with the final scope of evaluating the effect of the introduction of admission fees, as a new source of revenue (Santagata and Signorello, 2000; Maddison and Foster, 2003; Lampi and Orth, 2009). The second line of research looks at the economic impact of the museums and cultural initiatives in the promotion of local economic growth and development. It considers the indirect use value of museum and cultural attractions, their external effects on local tourism operators and their multiplier effect in the local economy (see, e.g., Cooke and Lazzaretti, 2008).

Available analyses, within both research lines, are generally based on case studies, and the conclusions are difficult to generalize (Bille and Schulze, 2006; Plaza, 2008). In the wake of the successful cases described, the 1990s saw an increase in the number of museums at the international level and, consequently, an increase in competition. Competition among local policy-makers also arose: they were confident that the setting of a museum could easily lead to increasing tourism flows, with the consequent economic growth of the local area.

However, sound quantitative evidence on this possible nexus is lacking: only few recent studies present econometric exercises on the relation between cultural tourism specialization and economic growth at national and regional level (respectively, Arezki et al., 2009, and Cellini and Torrisi, 2009). Similarly, only few contributions study the relation between the valorization of cultural attractions and tourist arrivals, from an econometric point of view; for instance, Yang et al. (2009) test the significance of the inclusion of monuments and sites in the UNESCO World Heritage List (and in national lists) in attracting international tourists to China. The results concerning the effectiveness of such cultural sites in attracting tourism are mixed.

No specific studies are available, to the best of our knowledge, on the causality between tourism flows and visit to monuments and museums. More explicitly, we think that it can be interesting to deal with the following question: is it the presence of cultural attraction (and specifically, museum and monuments) which attracts tourists, or –on the contrary– does the existence of tourism flows permit museums to be visited? We try to answer this question, taking aggregate Italian data into consideration.

3. THE ITALIAN CASE

We analyze Italian data with a monthly frequency over the period January 1996 to December 2007. Data are from ISTAT, the Italian Central Statistics Office, and they are easily obtainable from the ISTAT website (and from the website of the Ministry of Cultural Heritage in the case of time series of visits to museums and monuments).

As far as tourism variables are concerned, we consider tourist presences, measured by overnights (denoted by PRESTUR), tourist arrivals (ARRIV) and average stays (AVSTAY); as is well known, arrivals multiplied by average stays give the presences. Official data are articulated according to the source countries, the region of destination, the accommodation structure, and so on, but –when not differently stated– we refer to the total datum (the total presences, or total arrivals, and so on). Figure 1.a represents the pattern of the time series of overnights, while some descriptive statistics of such series are offered in Table 1 (line a). Arrivals and stays are described in panel b and c of Figure 1, and their statistical properties are summarized in Table 1 (lines b, c). Figure 1.d and Table 1 provide information related to the visits to State museums, monuments and museum networks. Also in this case, more articulated data are available, but generally we limit ourselves to the aggregate datum (MUSMONUV). Note that only cultural sites run by the State are considered here: though questionable, this is a necessary choice, due to the fact that consistent data are not completely available for monuments or museums run by private subjects or local public administrations; however, the main cultural sites are run by the State in Italy, and these museums account for over one third of the visits to museums (as documented, e.g., by Fedeli and Santoni, 2006), so we believe that our data are sufficiently representative.

INSERT FIGURE 1 AND TABLE 1

Clearly, monthly data show a great deal of variability and strong seasonal patterns. For this reason, we report some indices related to seasonality in the time series at hand (Table 2). Specifically, the Gini index provides information on the month concentration (the higher the Gini index, the stronger the seasonality concentration). Alternatively, one can decompose the time series (according to one of the available procedures) into trend-cycle, seasonal and erratic components, and take a look of the seasonal factors: the higher the variation field of the seasonal factors (or the higher their standard deviation), the more severe the seasonality. Again, one can take a look at the correlation between the original series and the seasonally adjusted series (the higher the correlation, the less important the seasonal component). The message from Table 2 is simple and clear: all the considered series have a significant seasonal component, even if seasonality in tourist presences appears to be more severe than the seasonality in museums and monuments attendance.

INSERT TABLE 2

Quite interestingly, the peak seasons, in tourism variables and in visits to museums and monuments, do not coincide: August is the peak season for tourism, while April represents the peak season for visits to museums and monuments; the same non-coincidence holds for the season with the lowest values, which is November for tourism flow variables and January for cultural sites' attendance. Apart from seasonality patterns, it is clear that arrivals show an upward trend, while the trend of average stay is decreasing; these facts are consistent with tourist presences which are rather stable in the long run. Visits to museums and monuments appear to have a slightly positive long-term tendency until 2005, and then a slight decreasing tendency emerges; so, they appear stable over the whole period sample.

We are interested in establishing which statistical representation is the most adequate for the data at hand. Not surprisingly, we will find that seasonal unit roots are present. This result is common to all recent applied analyses of tourism time series, in different countries and over different periods and frequencies (see, e.g., Lim and McAleer, 2000, 2001, Balaguer and Cantavella Jordà, 2002, Dritsakis, 2004, 2008, Brida, Carrera and Risso, 2008, on quarterly data referred to Australia, Spain, Greece and Mexico as destinations, respectively, and Goh and Law, 2002, Koc and Altinay, 2007, and de Olivera, 2009, on monthly data concerning arrivals at different destinations). Thus, we have to take a time series analysis approach based on seasonal unit root and seasonal integration and co-integration properties. We provide a methodological note on such techniques below. It is important to stress that the co-integration analysis will provide a natural way (and straightforward tools) to assess the direction of causality, which is the core point of interest in the present paper.

4. UNITS ROOTS AND CAUSALITY IN TIME SERIES ANALYSIS

4.1. A methodological note

The issue of unit-root has been introduced into statistics and economic analysis with reference to annual time series. As is well-known, a time series X_t is said to have a unit root, if in its autoregressive representation $X_t = bX_{t-1} + u_t$ (with t=1,2,...T), parameter b is equal to 1, and the error term u_t is a stationary process.

In order to detect the presence of a unit root in a time series, the procedure first suggested by Dickey and Fuller (1979) involves subtracting X_{t-1} from both sides of the autoregressive representation, so to obtain $\Delta X_t = cX_{t-1} + u_t$ (with $\Delta X_t \equiv X_t - X_{t-1}$ and $c \equiv b-1$). The presence of the unit root can be tested, by evaluating c=0 and by resorting to the specific critical values for the t-statistics in this case ((augmented) Dickey-Fuller test). If only one differentiation makes the series stationary, then the series is integrated of order one.

The statistical properties of an integrated series largely differ from the properties of a stationary series. In particular, an integrated series has no inherent tendency to return to mean value (that is, shocks on it have permanent effects) and it has increasing expected variance.

If two time series, each of whom integrated of order one, have a stationary linear combination, then they are said to be co-integrated. Loosely speaking, co-integration means that long-run relationship exists, as long as two co-integrated series cannot diverge "too much" from each other. The stationary linear combination can be interpreted as the long-run link between the non-stationary series. Operationally, in order to evaluate the presence of co-integration, a static equation is considered, say $Y_t = m + nX_t + ERR_t$ (t=1,2,...T). If the error term e_t is a stationary process, X and Y are co-integrated, and the residuals ERR_t can be interpreted as the "error" (or "discrepancy") of current variable *Y* with respect to its long-run equilibrium value dictated by the co-integrating relationship.

According to Granger's representation theorem (Granger, 1986, Engle and Granger, 1987), if two integrated variables co-integrate, an error correction mechanism is operative, which means that *Y* and/or *X* have to move in order to correct the disequilibrium with respect to the long-run relationship. This means that (at least) one Granger causal ordering does exist. Thus, the co-integration analysis offers powerful tools to look at the causality issue. Since we will use these concepts extensively, let us briefly summarize the idea behind the representation theorem. Consider the following system representing the dynamics of the co-integrated variables *X* and *Y*, where Δ is the first-difference operator and *ERR* denotes the error term of the static regression:

$$[1] \qquad \begin{aligned} \Delta Y_t &= a_Y + \gamma_Y ERR_{t-1} + \sum_h \varphi_{Yh} \Delta Y_{t-h} + \sum_k \phi_{Yk} \Delta X_{t-k} + e_{Yt} \\ \Delta X_t &= a_X + \gamma_X ERR_{t-1} + \sum_h \varphi_{Xh} \Delta Y_{t-h} + \sum_k \phi_{Xk} \Delta Y_{t-k} + e_{Xt} \end{aligned}$$

According to equations of system [1], variables *X* and *Y* move fro two reasons: (a) to adjust the long-run disequilibrium (that is,, in response to the term *ERR*) –this component is the error correction mechanism; and (b) in response to short-run variations of them (captured by the terms ΔY_{t-j} and ΔX_{t-j}).

The Granger representation theorem assures that at least one error correction mechanism exists if (and only if) two series are co-integrated. This means that parameter γ_Y and/or γ_X has to be significant (and negative) in at least one of the two equations of system [1].

The equations of system [1] with error correction mechanism allow us to define different concepts of causality. The long-run Granger-causality refers to the links between the *levels* of *Y* and *X*, and more precisely refers to the variable which has to move in order to adjust the "disequilibrium" with respect to the co-integrating relationship. Specifically, if γ_Y (or γ_X) is significant, it means that variable *Y* (or *X*)moves in order to reduce the disequilibrium with respect to the long-run equilibrium value; clearly, if only one error correction coefficient is significant, a one-directional causal link is established : if $\gamma_Y \neq 0$, variable Y is Granger-caused (in the long run) by X. If both γ_X and γ_Y differ from zero, bidirectional (long-run) Granger causality exists.

The short-run Granger-causality refers to the *differences* of Y and X. In each equations of [1], the components related to ΔY_{t-j} and ΔX_{t-j} are deemed to capture "short-run" determinants of ΔX and ΔY. If parameters $\{\phi_Y\}=0$ (which means that lagged values of ΔX_t do not affect contemporary value of ΔY_t) then ΔX_t does not Granger cause ΔY_t, or, X_t does not Granger-cause Y_t in its short run movements. Reversely, if $\{\varphi_X\}=0$, Y_t does not Granger cause X_t in the short-run components. (Different concepts of causality are reviewed by Granger, 1988).

Different techniques are available to measure the strength of causal links. For instance, Pesaran and Shin (2002) suggested the variance decomposition technique: the variable whose variance is explained by its own past value in the largest part, is the "most exogenous" one. Granger and Lin (1995), in the framework of co-integration, proposed to measure the strength of causality of Y on X by means of the following index:

[2] $$M_{Y\to X} = \log\left[1+\frac{{\gamma_X}^2(1-C^2)}{(\gamma_X C-\gamma_Y)^2}\right] \quad ,C = corr(e_X, e_Y)$$

Clearly, if γ_X is not different from zero, then $M_{Y\to X}=0$, i.e., Y does not cause X. This kind of techniques have been extensively used in applied macroeconomic analysis (especially during the 1990s), using annual data.

The extension of the integration / co-integration analysis to seasonal series can be dated back to Dickey, Hasza and Fuller (1984). Fransen (1996) or Ghysels and Osborn (2001) offer comprehensive reviews of theoretical aspects and applied investigations of seasonal integration and co-integration.

According to standard definition (see, e.g., Ghysels and Osborn, 2001, Def. 3.1) the non-stationary stochastic process Y_t, observed at s equally spaced time interval, is said to be seasonally integrated of order one if $\Delta_s Y_t = Y_t - Y_{t-s}$ is stationary. The symbol Δ_s –often called the "seasonal differencing filter"– denotes the first-difference of lag s (in monthly data, s=12). In other words, $\Delta_s Y_t$ denotes the difference of the realization in any given season with respect to the realization of the variable in the same season of the previous year. The simplest case of a seasonally integrated process is a season random walk, which is described by the data generating process $\Delta_s Y_t = \varepsilon_t$, that is, $Y_t = Y_{t-s} + \varepsilon_t$, with ε_t denoting a white noise process. More generally, seasonally integrated processes can possess drift(s), i.e., a constant term, or different constant terms for different seasons; they can possess a deterministic trend or a stationary ARMA structure of the error term. For the specific purpose of the data at hand, we consider monthly data, and we will consider an equation of type

$$[3] \qquad Y_t = a + \rho Y_{t-12} + v_t$$

or, subtracting $Y_{t\text{-}12}$ from both sides,

$$[4] \qquad \Delta_{12} Y_t = a + \alpha Y_{t-12} + v_t$$

with $\alpha = \rho - 1$.We are interested in evaluating whether $\rho = 1$, i.e., $\alpha = 0$; if such a hypothesis is accepted (rejected), the series is "seasonally integrated" ("seasonally stationary").

Prior to the decision about seasonal stationarity, however, we have to take decisions about three different points. Firstly, we have to evaluate whether 12 different constant terms are appropriate (one for each season) instead of one constant term; in such a case, a has to be interpreted as $a = \{a_i\}_{i=1}^{12}$. Operationally, we evaluate whether 11 additional seasonal dummy variables beyond a constant are significant (see also Fransen and Kunst, 1999 on this point); generally, the inclusion of seasonal dummies turns out to be appropriate in our present cases. Secondly, we have to evaluate if a deterministic trend (T) is appropriate. Generally, the deterministic trend is significant in our data; the inclusion of a trend makes the test less powerful, but we anticipate that our conclusions are robust to the omission of the time trend. Thirdly, we evaluate whether to introduce a number of autoregressive terms of $\Delta_{12} Y_t$ in order to have white noise regression residuals; in most cases, the 1st, 2nd and 12th lags of the dependent variable are statistically significant and sufficient to make white noise residuals, and hence they are inserted in the regression.

In sum, the following regression equation is considered in the applied analysis:

[5] $$\Delta_{12}Y_t = \sum_{i=1}^{12} a_i + \tau T + \alpha Y_{t-12} + \sum_j \beta_j \Delta_{12} Y_{t-j} + \varepsilon_t$$

For testing for the presence of the seasonal unit root, we look at the significance of the coefficient α. Also in this case, the distribution of estimated standard error, and the Student–*t* statistics, are non standard, and specific tabulations of critical values are necessary. The tabulation of critical values is provided by Dickey, Hasza and Fuller (1984). Different Tables are appropriate, depending on whether no-constant, or a unique constant or different *s* constant terms are introduced. Dickey, Hasza and Fuller label these models as "zero-mean model", "one mean model", and "seasonal means model", respectively.

If the null of seasonal unit root is not rejected (i.e., $\alpha = 0$ or equivalently $\rho = 1$), the series is seasonally integrated. The substantial meaning of such a conclusion is very important. Seasonally integrated series possess *s* unit root processes (one for each of the *s* seasons), none of which has a tendency to return to a deterministic path.

Two seasonally integrated time series X_t and Y_t are seasonally co-integrated, if a linear combination exists which is seasonally stationary. Operationally this means that the residuals from a regression involving X_t and Y_t (and possibly other deterministic components, like time trend and seasonal dummies) have to be seasonally stationary. In concrete terms, we have to run a regression (called "static co-integrating regression") of Y_t on X_t (or X_t on Y_t), and then we consider the regression residuals and perform the seasonal integration tests on them: if the null hypothesis of seasonal integration in the residuals is rejected, then *X* and *Y* are seasonally co-integrated. This test perfectly corresponds to the Augmented Dickey Fuller test, and critical values are provided by Dickey, Hasza and Fuller, as already mentioned.

In concrete terms, provided that X_t and Y_t are seasonally integrated, we will run the (static) regression

[6] $$Y_t = \sum_{i=1}^{12} a_i + \tau T + mX_t + e_{1,t}$$

from which we save the fitted series $e_{1,t}$; in order to evaluate its seasonal stationarity, we run a regression similar to [4], and specifically:

[7] $$\Delta_{12} e_{1tt} = \alpha e_{1,t-12} + \varepsilon_t$$

(possibly augmented by lagged terms of $\Delta_{12}e_{1,t}$, to render residual ε_t stationary, but without the constant term, the mean of regression residuals being zero) and we look at the Student-*t* of coefficient α. If stationary, the series $e_{1,t}$ can be interpreted as the linear combination which represents the "error" or the discrepancy with respect to the co-integrating relationship. (We also will consider the regression of *X* on *Y*, and perform the same test of seasonal stationarity on the residuals from this equation ($e_{2,t}$). The conclusions about co-integration of time series have to coincide –and this happens, in fact, in all the cases considered below).

Also in this case, the Granger representation theorem can apply: if two seasonally integrated series co-integrate, at least one error-correction mechanism is operative, and the causal link can be detected, in the sense that it is possible to establish if *X* or *Y* (or both) move over time to correct the discrepancy with respect to the co-integrating relationship equilibrium values. If the series co-integrate, the error correction mechanism has to be operative, according to the lines of the Granger representation theorem. Thus, we will consider the two equations :

$$\begin{aligned} \Delta_{12}Y_t &= a_Y + \gamma_Y ERR_{1,t-12} + \sum_h \vartheta_{Yh}\Delta Y_{t-h} + \sum_k \theta_{Yk}\Delta X_{t-k} + e_{Yt} \\ \Delta_{12}X_t &= a_X + \gamma_X ERR_{2,t-12} + \sum_h \vartheta_{Xh}\Delta Y_{t-h} + \sum_k \theta_{Xk}\Delta X_{t-k} + e_{Xt} \end{aligned} \tag{8}$$

and we will look at the coefficients γ_Y, γ_X to derive conclusions about the long-run Granger causality, and at the coefficients $\{\vartheta_{Yh}\}, \{\theta_{Yk}\}$ to study the short-run causality.

4.2 Evidence: tourist presences, and the attendance at monuments and museums

We aim to test the presence of seasonal unit root in the time series of tourists' presence and cultural sites' attendance in Italy. Relevant regression results are reported in Table 3. Specifically, we report the results for equations specified as in [5]. In all the cases at hand, we find what follows. First, the introduction of different seasonal dummies is appropriate (see Column (2): test *F* on the significance of additional 11 dummies beyond a constant term always leads to the conclusion that different additional dummies are different from zero). Second, a deterministic time trend is significant, and it is inserted; however, since the power of unit root test is low when a time trend is inserted, we preferred to check also the results from the specification without the deterministic trend: the conclusion about the presence of the seasonal unit root was the same, so that we can conclude that our results are robust to the choice of including the deterministic time trend or not. Third, different numbers of lags in the short-term dynamics of the equation are appropriate: with reference to specification [5], j may vary across different equations; however, for the series at hand, the significant lags are 2 and 12. Last but not least, the presence of the seasonal unit root at periodicity 12 cannot be rejected: the Student-t statistics of the estimated α is -3.90 and -3.68 for PRESTUR and MUSMONUV, respectively. These figures are smaller –in absolute value– than the critical level -5.86, tabulated by Dickey, Hasza and Fuller for the usual 95% significance level; (in the absence of the deterministic trend, the Student-t of estimated alpha, would be -2.72 and -3.68, respectively, leading to the same conclusion of seasonal integration.)

INSERT TABLE 3

Thus, the conclusion that the series at hand are seasonally integrated is out of any doubt. The same conclusion holds for alternative –though not advisable– specifications of the regression equation, considering alternative design of the deterministic components, like time trend which assume one value for each year. It is interesting, hence, to establish whether co-integration links exist. In advance, it is advisable to take a look at the pattern of the two series in Figure 2: panel (a) shows the raw data, panel (b) normalizes the data to have the same adjusted mean; panel (c) provides the scatter-plot, and the existence of different seasons is clear.

INSERT FIGURE 2

In order to establish the possible existence of co-integration, we consider the relationships corresponding to [6], with tourist presence or, alternatively, the cultural site attendance as the dependent variable; we save the regression residuals and perform the seasonal stationarity test on them. Table 4 shows that, in any case, the regression residuals are stationary, leading to the conclusion that tourists presence and cultural sites attendance co-integrate. This result is also consistent with the evidence coming from the estimation of the equations containing the error correction mechanism. In particular, we estimate system [7] with museum and monuments attendance in the place of Y and tourist presences in the place of X. The results are provided in Table 5

INSERT TABLE 4 AND TABLE 5

It is clear that the error correction term is significant in both equations –the equation explaining the visits to cultural sites, and the equation concerning tourist presence– leading to the conclusion that bi-directional long-run causal links are present. In other words, we find that the tourist presences Granger-cause the visits to museums and monuments, and the visits to museums and monuments Granger-cause the tourist presences, in the long run. However, the coefficient is larger in absolute value in the case of visits to cultural sites, suggesting that this variable is more reactive to long-run disequilibria.

As to the short-run Granger-causality, the conclusion is sharper, in the sense that the causality links run from (variation of) presence to (variation of) visits to cultural sites. This is clear from the Student-t statistics of single coefficient (in Table 5), and can be confirmed by F type tests on the significance of multiple coefficients: a test on the significance of lags 1 and 12 of D12PRESTUR in the equation of D12MUSMONUV provides F=3.47 (p=0.034), while a test on the significance of lags 1, 2, and 12 of D12MUSMONUV in the equation of D12PRESTUR gives F=0.86 (p=0.460).

As already mentioned, for discerning the endogenous/exogenous nature of variables, including in the context of co-integration analysis, some authors apply the "generalized variance decomposition" technique (Pesaran and Shin, 2002 ; see also Masih et al., 2009 for a very recent application): the relative exogeneity or endogeneity of a variable can be detected by the proportion of the variance explained by its own past. The variable which is explained mostly by its own shocks is the «most exogenous». The conclusion in the present case (see Table 6) is very clear: it is tourist presence that leads (rather than lags) visits to cultural sites. The same conclusion emerges, based on the computation of the Granger Lin causality strength index: the strength of causality from tourist presence to site visits is log(1.50), while the strength of causality from visits to presence is log(1.37).

INSERT TABLE 6

4.3 Tourist arrivals and stays

We can continue to employ the co-integration analysis approach to investigate the links of tourist arrivals and average stays, on the one side, with the visits to cultural sites on the other side (Italy, January 1996 – December 2007). This makes sense since the monthly time series of arrivals and stays –like tourist presences– possess a seasonal unit root at periodicity 12, as shown in Table 7 (where we report only the estimates of the specification with the deterministic trend; however, the conclusion on the presence of the seasonal unit root is the same, even if we omit the deterministic trend); other information on the time series were already provided in Tables 1 and 2.

INSERT TABLE 7

Even in the cases of arrivals and stays, we find that each of such tourism series co-integrates with the series of cultural sites' visits, as documented by Table 8.

The Table reports the results from the static co-integrating regression (along the lines of equation specification [6]) and the results from the dynamic specification like [8]. In the static regression, twelve seasonal dummies are introduced, since additional dummies for seasons are significant; in the dynamic equation for evaluating the error correction mechanism, an appropriate number of lags of the dependent variable is introduced, following a specification strategy from the general to the particular, which started considering the lags of order, 1,2, 3, 4, 12, 24 and maintained only the significant ones (95% significance level).

INSERT TABLE 8

The evidence concerning the causality links over the long run is very clear. Arrivals and visits to museums and monuments Granger-cause each other; however, like in the case of presences, the quantitative dimension of the error correction coefficient suggests that cultural visits adjust to arrivals in a larger extent than the reverse. As far as stays are concerned, one can see that stays Granger-cause one-directionally visits to monuments and museums in the long run (the error correction is significant only in the equation explaining the dynamics of cultural site visits). The Granger and Lin causality index lead to the same substantial conclusion: the stronger causal link goes from arrivals and stays to visits to museums and monuments; a similar conclusion is provided by the variance decomposition technique, which suggest that arrivals and stays are "more exogenous". In the short-run dynamics, arrivals and visits to cultural sites cause each other, while a one-directional link emerges as far as stays are concerned: stays do not cause visits, while visits cause stays.

These pieces of evidence lend themselves to some considerations. Loosely speaking, the long-run dynamics have to do with the long-term decisions of people. It is the dynamics of visits to cultural sites that adjust to the dynamics of tourism flows. Thus, it is hard to sustain that cultural site visits play a long-run promoting role with respect to tourism flows: provocatively stated, it is false that tourists plan to come and stay in Italy in order to visit cultural sites; rather, people visit museums and monuments just because they decided to arrive and stay in Italy. However, in the short run, some significant causal effect of the visit of cultural sites emerges upon the average stays. Just to give a simple and intuitive explanation, imagine that people have planned the holiday; the presence of cultural sites has been ineffective at that stage; however, if the weather is bad (short-term shock), the presence of cultural attractions can be effective in convincing people to remain rather than to go home in advance. More seriously, the presence of cultural attractions is ineffective in determining long-run dynamics, but can be effective in the short-run decisions of people.

We are in a position now to provide estimates of the elasticity of cultural site visits to tourist variables. These elasticities are shown in Table 9, which considers both the unconditional elasticity estimates, and the estimates from the model with multiple seasonal dummies. The values, however, are rather similar. All estimates are statistically significant. Elasticity of visits to museums and monuments with respect to tourist presence is around 0.86; a test of equality of such a value to 1 rejects this hypothesis; the elasticity with respect to arrivals is around 0.9: also in this case such a value turns out to be statistically different from 1. Elasticity with respect to stay is about 10: a 1% increase in average stay entails a 10% increase in visits to museums and monuments.

INSERT TABLE 9

5. DISCUSSION, IMPLICATIONS AND CONCLUSIONS

The effectiveness of cultural attractions in enhancing tourism is a point of interest not only for academics, but also for private subjects and policy-makers. In several cases, the presence of cultural attractions is deemed to act as an engine for attracting tourism flows, or qualifying the tourism. Nevertheless, the empirical evidence on the relationships between cultural attraction attendance and tourism flows is limited to interesting but specific cases. The present paper has aimed to fill this absence, providing an analysis on aggregate data.

We have taken Italy as a case study, and have analyzed monthly data over more than a decade, referring to tourist overnights, arrivals, and average stays, on the one hand, and visits to museums and monuments on the other. We have proved that all series possess a seasonal unit root, i.e., they are seasonally integrated. Strong evidence of co-integration between tourism flows and museum and monument attendance has emerged. This means that long-term relationships exist between these variables. More importantly, Granger causality analysis has permitted to conclude that a one-way direction of causality generally emerges, and tourism flows Granger-cause the attendance to cultural sites, in the long run.

The conclusion about the causality nexus, running from tourism flows to museum attendance, is the core result of the present research, and lends itself to two comments. First, from a substantial point of view, we can state that museums cannot be requested, on average, to play a role as major tourism attractors. The available literature on specific successful cases (generally superstar museums, which represent a minority among museums) has perhaps generated the misleading idea that museums can be primary engines for tourism and hence for growth. We rather believe that, in general, museums can be the "icing on the cake" in a destination in which a bundle of several material and immaterial structures are the roots of tourism attraction: museum and monument visits could be able to determine longer average stays, rather than larger arrivals. Second, from a methodological point of view (even if one has to be aware that museums and monuments play roles that go well beyond tourism attraction), omitting the tourism variables from the set of the determinants of the attendance at museums and monuments can be seriously misleading, in the presence of documented co-integrating relationships.

Other conclusions are possible, concerning the role of cultural attractions as a means to reduce seasonality in tourism flows. Schematically, the idea could be as follows: provided that the visits to cultural sites show a lower degree of "overall" seasonality than tourism arrivals or presences (as documented also in our Table 2 for the Italian case), and provided that peaks of cultural visits are in spring months, rather than in summer, the promotion of cultural tourism should help in reducing tourism seasonality and hence congestion. Unfortunately, our analysis clearly shows that cultural visits follow, rather than lead, tourism presences and arrivals. Provocatively, visits to cultural sites are perceived by most tourists as a by-product of a holiday stay, rather than the main goal. Consistently, cultural heritage attractions seem to be effective tools to differentiate tourism products; their effectiveness in reducing tourism seasonality appears to be more questionable.

REFERENCES

Algieri, B. (2006). An econometric estimation of the demand for tourism: the case of Russia. *Tourism Economics*, 12, 5-20.

Alleyene, D. (2006). Can season unit root testing improve the forecasting accuracy of tourist arrivals?. *Tourism Economics*, 12, 45-64.

Arezki R., R. Cherif and J. Piotrowski (2009). Tourism specialization and economic development: Evidence from the UNESCO World Heritage List. IMF working paper 09/176.

Balaguer, J. and Cantavella-Jordà, M. (2002). Tourism as a Long-run Economic Growth Factor: the Spanish Case. *Applied Economics*, 34, 877-884.

Bille, T. and G. G. Schulze (2008). Culture in urban and regional development. In V. A. Ginsburgh and D. Throsby (Eds.), *Handbook of the Economics of Art and Culture*, Amsterdam: North Holland - Elsevier, 1051-1099.

Brida J. G., E. Carrera and W.A. Risso (2008). Tourism's Impact on Long-Run Mexican Economic Growth. *Economics Bulletin*, 3(21), 1-8.

Cellini, R. and G. Torrisi (2009). The regional public spending for tourism in Italy: An empirical analysis. MPRA working paper n. 16131

Cooke, P. and L. Lazzaretti (2008). *Creative Cities, cultural clusters and local economic development*, Cheltenham: Edward Elgar.

de Olivera Santos G.E. (2009). Forecasting tourism demand by disaggregating time series: Empirical evidence from Spain. *Tourism Economics*, 15, 467-72.

Dickey D.A., Hasza D. P., and Fuller W. A. (1984). Testing for Unit Roots in Seasonal Time Series. *Journal of the American Statistical Association*, 79, 355-367.

Dritsakis, N. (2004). Tourism as a Long-run Economic Growth Factor: An Empirical Investigation for Greece Using Causality Analysis. *Tourism Economics,* 10, 305-316.

Dritsakis, N. (2008). Seasonal Analysis of Tourist Revenues: An Empirical Research for Greece. *Tourismos*, 3(2), 57-70.

Engle R. F. and Granger C. (1987). Co-integration and Error-Correction: Representation, Estimation and Testing. *Econometrica*, 55, 251-276.

Engle R. F., Granger C. W. J, and Hallman J.J. (1989) "Merging Short- and Long- Run Forecasts. An Application of Seasonal Cointegration to Monthly Electricity sales Forecasting. *Journal of Econometrics*, 40, 45-62.

Engle R. F., Granger C. W. J, Hylleberg S., and Lee H. S. (1993). Seasonal Cointegration: the Japanese Consumption Function. *Journal of Econometrics*, 55, 275-303.

Fedeli S., and M. Santoni (2006) "The Government's Choice of Bureaucratic organization: An Application to Italian State Museums. *Journal of Cultural Economics*, 30, 41-72.

Fransen, P.H. (2006). Recent advances in modeling seasonality. *Journal of Economic Survey*, 10,299-345.

Fransen P.H. Kunst R.M. (1999). On the Role of Seasonal Intercepts in Seasonal Cointegration. *Oxford Bulletin of Economics and Statistics*, 61, 409-433.

Frey B. S. and S. Meier (2006). The Economics of Museums. In V. A. Ginsburgh and D. Throsby (Eds.), *Handbook of the Economics of Art and Culture*, Amsterdam: North Holland - Elsevier, 1017-1042.

Goh, C., and R. Law (2002). Modelling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, 23: 499-510.

Granger, C. W. J. (1988). Some recent developments in a concept of causality, *Journal of Econometrics*, 39, 199-211.

Granger C. W .J. and Lin J.-L. (1995). Causality in the Long Run. *Econometric Theory*, 11, 530-36.

Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2, 111-120.

Gregory A. W. and Hansen B. E. (1996) "Residual-Based Tests for Cointegration in Models with Regime Shifts. *Journal of Econometrics*, 70, 99-126.

Ghysels E., Osborn D. R. (2001). *The Econometric Analysis of Seasonal Time Series*, Cambridge: Cambridge University Press.

Hylleberg S. (1995). Tests for Seasonal Unit Roots. General to Specific or Specific to General?. *Journal of Econometrics*, 69, 5-25.

Inder B., "Estimating Long-run Relationships in Economics. *Journal of Econometrics*, 75, 53-68.

Kremers J., N. Ericsson, and J. Dolado (1992) "The Power of Cointegration Tests. *Oxford Bullettin of Economics and Statistics*, 54, 325-348.

Koc E., G. Altinay (2007) , "An analysis of seasonality in monthly per person tourist spending in Turkish inbound tourism from a market segmentation perspective. *Tourism Management*, 28, 227-237.

Kotler N. and P. Kotler (1998). *Museum strategy and marketing*. San Francisco: Jossey-Bass Inc Publisher

Kunst R. M. (1997). Testing for Cyclical Nonstationarity in Autoregressive Processes. *Journal of Time Series Analysis*, 18, 123-135.

Johansen S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, 231-54.

Johansen S. and C. Juselius (1990).”Maximum likelihood estimation and inference on cointegration – with application to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

Johnson P. S. (2003). Museum. In R. Towse (Ed.), *A handbook of cultural economics*, Cheltenham: Edward Elgar, 315-320.

Lampi, E. and M. Orth (2009). Who visits the museums? A comparison between stated preferences and observed effects of entrance fees. *Kyklos*, 62, 85-102.

Lim C., and M. McAleer (2000). A Seasonal Analysis of Asian Tourist Arrivals to Australia. *Applied Economics*, 32, 499-509.

Lim C., and M. McAleer (2001). Cointegration Analysis of Quarterly Tourism Demand by Hong Kong and Singapore for Australia” *Applied Economics*, 33, 1599-1619

Maddison, D. and T. Foster (2003). Valuing congestion costs in the British Museum. *Oxford Economic Papers*, 55, 173-190.

Masih M., A. Al-Elg A., and H. Madani (2009). Causality between Financial Development and Economic Growth: An Application of Vector Error Correction and Variance Decomposition Methods to Saudi Arabia. *Applied Economics*, 41, 1691-99.

Narayan P. K. (2004) “Fiji tourism demand: the ARDL approach to cointegration. *Tourism Economics*, 10, 193-206.

Phillips P. C. B. and M. Loretan (1991) “Estimating Long-Run Economic Equilibria. *Review of Economic Studies*, 58, 405-35.

Plaza, B. (2008). On some chellenges and conditions for the Guggenheim museum Bilbao to be an effective economic re-activator. *International Journal of Urban and Regional Research*, 32, 506-517.

Santagata, W. and G. Signorello, (2000). Contingent valuation and cultural policy: the case of Napoli Musei Aperti. *Journal of Cultural Economics*, 24, 181-204.

Yang, C.H., H.L. Lin. and C.C. Han (2009). Analysis of international tourist arrivals in China: The role of World Heritage sites. *Tourism Management*, in press.

TABLES

Table 1 – Descriptive statistics on variables

Variable	Average	Min – Max (date)- (date)	St. Dev.
a) PRESTUR (million)	27.850557	8.529030-78.026590 (Nov. 1997)-(Aug. 2007)	19.481268
b) ARRIV (million)	6.773693	3.141226-13.110528 (Nov. 1997)-(Aug. 2007)	2.709173
c) AVSTAY (days)	3.762	2.670-6.619 (Nov. 1998)-(Aug. 2001)	1.118
d) MUSMONUV (million)	2.505890	0.770116-4.598806 (Jan. 1997)-(Apr. 2006)	1.002944

Table 2 – Descriptive statistics on seasonality in data

Variable	Gini index	Seasonal Factors: St. Dev.	Seasonal Factors: Min - Max (date)- (date)	Corr (Raw Series, Seas-Adjust Series)
PRESTUR	0.367	0.693	0.356-2.710 (Nov. 1996)-(Aug. 1996)	0.124
ARRIV	0.227	0.385	0.542-1.764 (Nov. 1997)-(Aug. 1996)	0.257
AVSTAY	0.154	0.296	0.714-1.696 (Nov. 1996)-(Aug. 1997)	0.089
MUSMONUV	0.254	0.385	0.415-1.774 (Jan. 2002)-(Apr. 1996)	0.300

Notes: Column (1) reports the Gini index on monthly data of the raw series. Columns(2) to (4) take into account the seasonal adjustment computed with Census X-12-Arima adjustment programme: Column (2) and (3) report the standard deviation and the Min-Max values of the seasonal factors, while Column (4) reports the correlation between the original and the seasonally adjusted series.

Table 3 – Seasonal unit root – Regression Equation [5]

	(1) Estimated coefficient α (t)	(2) Seasonal dummies	(3) Determ time trend	(4) Lags of the dependent variable	(5) R2 DW
PRESTUR	-0.25 (-3.90)	YES $F_{11,104}$=4.03 [p=0.000]	13791.7 (2.68)	2;12	0.41 1.94
MUSMONUV	-0.21 (-3.68)	YES $F_{11,118}$=4.51 [p=0.000]	210.4 (2.71)	2;12	0.41 1.82

Notes: Column (1) reports the estimate of coefficient α in specification [5] and its Student-*t* statistics. The value of Student-*t* statistics has to be compared to the critical values reported in Table 5 or 7 in Dickey-Hasza and Fuller (1984); critical value is -5,86 in the case where different seasonal dummies are inserted in the regression. Column (2) states whether seasonal dummies are introduced, and presents a F test (and its p-value in squared brackets) on the significance of additional 11 dummies added to the constant term. Column (3) reports the estimate of the deterministic trend coefficient, if inserted. Column (4) lists the lags of the dependent variable inserted in the regression to render residuals white noise: we started by considering lags 1,2,3,4,12 and decide to insert only the significant lags. Column (5) reports the R-squared and the DW statistics.

Table 4- Unit root test on the cointegrating regression residuals

	Residuals from : Visits on Presences	Residuals from : Presences on Visits
Estimated coefficient α (Student-t)	-0.64 (-8.11)	-0.59 (-7.61)
Lags of Dependent variables to have white noise errors	1; 2.	1; 2.

Notes: the 12th difference of the fitted residuals from the static regression equations is regressed against the 12th lag of the residual levels, according to eq. [10]. No constant term is inserted. Critical value at the 95% significance level for the Student-*t* is -1.77 (Dickey, Hasza, Fuller, 1984, Table 3).

Table 5 – Models with error correction mechanism – Estimation of Equations [8]

	Dependent variable : D12MUSMONUV	Dependent variable : D12PRESTUR
CONSTANT	84328 (4.51)	401518 (3.22)
ERR(-12)	-0.33 (-3.17)	-0.29 (-4.67)
D12MUSMONUV(-1)	*Ns*	*Ns*
D12MUSMONUV(-2)	0.20 (2.58)	*Ns*
D12MUSMONUV(-12)	-0.26 (-2.72)	*Ns*
D12PRESTUR(-1)	0.02 (1.90)	*Ns*
D12PRESTUR(-2)	*Ns*	0.38 (5.08)
D12PRESTUR(-12)	-0.03 (-2.83)	*Ns*
R2	0.379	0.315
F	14.04 [p=0.0000]	29.17 [p=0.0000]
Residuals autocorrelation: DW	1.63	1.85
Residuals autocorrelation: F test	F=1.34 [p=0.257]	F=0.37 [p=0.572]

Notes: *Ns* denotes "non-significant" (and hence the regressor is omitted from the chosen specification); Student-*t* statistics in parenthesis; autocorrelation F test is Breusch- Godfrey Serial Correlation LM test, with 4 lags.

Table 6 – Variance decomposition analysis

	TOURIST PRESENCE	MUSEUM ATTENDANCE
Lags 1,2	24.6%	11.3%
Lags 1,2,3,4	26.3%	11.7%
Lags 1,2,3,4,12	33.5%	32.4%

Notes: The Table reports the percentage of variance of Δ_{12} PRESTUR and Δ_{12} MUSMONUV explained by their own lagged values (a regression model is considered with seasonal dummies, but the conclusions do not change in the present of a single constant or no constant)

Table 7- Seasonal Unit root in the series of tourist arrivals and average stays

	(1) Estimated coefficinet α (*t*)	(2) Seasonal dummies	(3) Determin time trend	(4) Lags of the dependent variable	R2 DW
ARRIV	-0.21 (-3.21)	YES $F_{11,104}$=2.86 [p=0.002]	45856.6 (3.26)	2;	0.23 2.38
AVSTAY	-0.03 (-1.41)	YES $F_{11,104}$=3.36 [p=0.001]	-0.003 (-1.37)	2;	0.12 2.01

Notes: Columns are like in Table 3.

Table 8 – Cointegration Analysis for Tourist Arrivals and Stays with Cultural sites' visits

	X : Arrivals Y : Cultural sites' visits	X : Stays Y : Cultural sites' visits
Regression of X on Y (i.e., X=*f*(Y))		
Static regression results		
• Unit root in residuals: ADF (*t*)	(-8.42)	(-7.27)
Dynamic regression with ECM		
• EC Coefficient (and its *t*)	-0.22 (-2.38)	-0.03 (-1.20)
• Significant lags of ΔX	2	2;12
• Significant lags of ΔY	1	1;2
Regression of Y on X(i.e., Y=*f*(X))		
Static regression results		
• Unit root in residuals (*t*)	(-9.01)	(-2.85)
Dynamic regression with ECM		
• EC Coefficient (and its *t*)	-0.58 (-4.36)	-0.28 (-4.71)
• Significant lags of ΔX	1;2;12	Ns
• Significant lags of ΔY	2;12	2;12

Table 9- Estimates of Elasticity of cultural sites' visit (*Y*) w.r.t. tourism variables (*X*)

	Unconditional Elasticity	Elasticity conditional on seasonal means
$E_{Y,X}$: X=presences	0.864	0.855
$E_{Y,X}$: X=arrivals	0.935	0.924
$E_{Y,X}$: X=stays	10.93	10.98

Notes: Column (1) reports the estimate of the coefficient of regressor ln(X), against regressand ln(Y); in Column(2) seasonal dummies are inserted as additional regressors.

FIGURES

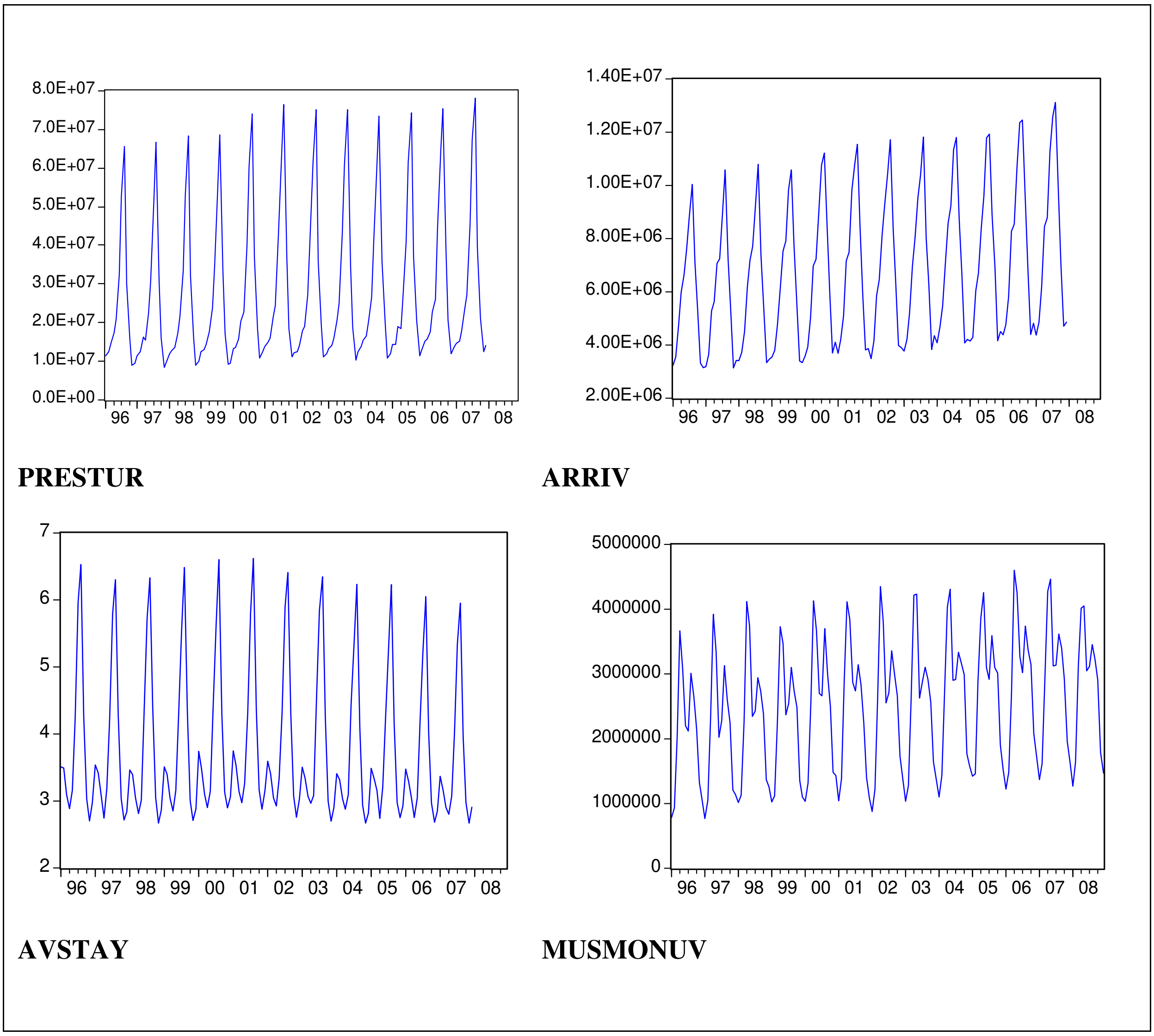

Figure 1 – Patterns of variables

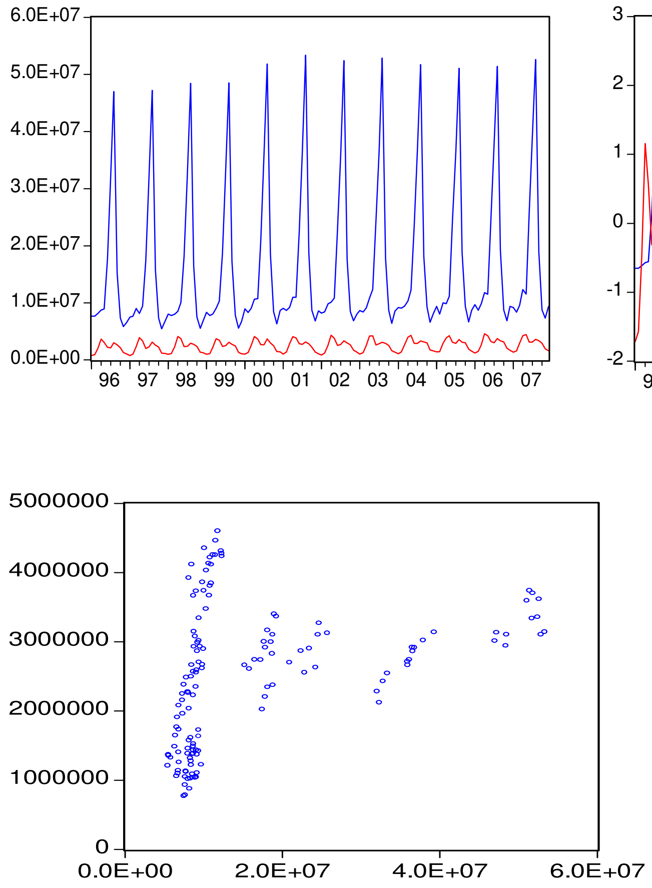

Figure 2 – Tourist Presences and Attendance to Museums and Monuments (patterns, normalized patterns, and scatter)