Migration and Tourist Flows

Nuno, Carlos Leitão and Muhammad, Shahbaz

ESGTS, Polytechnic Institute of Santarém and CEFAGE- University of Évora, Portugal, COMSATS Institute of Information of Information Technology, Lahore, Pakistan

9 November 2011

Online at https://mpra.ub.uni-muenchen.de/34905/
MPRA Paper No. 34905, posted 21 Nov 2011 08:52 UTC
Abstract

This study considers the relationship between immigration and Portuguese tourism demand for the period 1995-2008, using a dynamic panel data approach. The findings indicate that Portuguese tourism increased significantly during the period in accordance with the values expected for a developed country. The regression results show that income, shock of immigration, population, and geographical distance between Portugal and countries of origin are the main determinants of Portuguese tourism.

Key words: Tourism demand; panel data; immigration and Portugal.

JEL: C23, L83, M21
1. Introduction

When the economic geography was revisited in the 1990s some authors such as Krugman (1991) explained the relationship between North and South by introducing mobility between regions. This process involves the phenomenon of migration. Portugal is located in the South of Europe and we are seeing an increase in immigration. The process of immigration in Portugal is relatively recent.

With globalization, many nations have liberalized their trade policies and trade barriers removed. The relationship between tourism and immigration in Portugal has not been investigated till yet. Immigration and tourism have a complex and dynamic nature. One reason for this descriptive relation is associated with the reasons for the visit, in the case of immigration (foreign population in the country) involves family and friends (Jackson, 2003 and Yuan et al. 1995). The studies by Dwyer et al (1992), Hollanger (1982), Oigenblick and Kirschenbaum (2002), Fischer (2007) and Seetaram (2008) showed that immigration promotes tourism. In other words, immigrants are key drivers of the host country.

Both immigration and tourism have increased in recent decades. According to United Nation (2002) international migration in the world has increased from 154 million in 1990 to 175 million in 2000.

The per capita income, price index, geographical distance and their cultural surroundings (borders) are other explanatory variables, commonly used with greater frequency as determinants of tourism demand. The empirical studies (Gray, 1982, Kulendran and Wilson 2000, 2007 Dougan, Phakdisoth and Kim 2007) indicated that trade and tourism are positively correlated.

Trade and the phenomenon of migration is a major player in the international economics. The study of tourism and migration helps to explain the mobility of the population. Migration flows are associated in large measure to the improvement of living conditions. Either tourism or migration is associated with the location advantages (Illes 2004, Michalkóe Váradi 2003).

The historical and cultural proximity reduces transaction costs and promotes the phenomenon of migration.

Kirschenbaum (2002) pointed out that immigration has a positive impact on tourism demand in host countries.

This paper uses an unbalance panel data for international tourist flows to Portugal, from 16 countries, for the period 1995-2008. The structure of the paper is as follows. Section 2 presents the literature review and empirical studies. In section 3, we present the hypothesis. Section 4 shows the methodology. Section 5 presents the econometric model. The final section provides conclusions.

2. Immigration

According to the European Union (2007) and the Regulation no. 862/2007 of the European Parliament, the concept of immigration involves the movement to a host country for a period exceeding one year. The phenomenon of immigration in Portugal can be explained in four stages. The first phase emerges after the colonial period between 1975 and 1985. The Portuguese, from the former colonies, returned to their homeland, but also watched the arrival of African immigrants in Cape Verde, Guinea-Bissau and Angola. Cape Verde migratory movements had already begun in the 1960s. The second phase occurred in 1986 and ended in the 1990s. This phase is characterized by cultural and linguistic proximity, i.e., Brazilian immigrants. The third migratory phase emerges in the late 1990’s and ends at the beginning of the millennium, characterized by the immigration of Brazilian origin and the immigration from
Eastern European countries. A fourth phase started in the early years of the current decade in which we can mention the following: i) the period of economic crisis, in which we observe a decrease of immigration from Eastern Europe; ii) the balance from Portuguese speaking countries (PALOPS); iii) new entry of Brazilian immigrants.

In Figure 1, the main foreign residents in Portugal are displayed. The Brazilian, Cape Verdean and Ukrainian residents are the ones that stand out, followed by Romanians.

![Figure 1](source: Portuguese Border Services and authors’ calculation)

**Figure 1** Foreign Residents Major Nationalities (2000)

3. **Literature Review and Empirical Studies**

The studies of tourism and migration have been developed independently of one another up to second half of the twentieth century (Bell
Tourism as a form of temporary international migration can, like other types of movement, shift in the distribution of population.

Tourist visits can take place for various reasons: holidays, business trips, visit to friends and relatives, and others.

O'Reilly (1995) evaluated the relationship between migration and tourism. The author identified five categories: expatriates, residents, seasonal visitors, migrants and tourists.

Williams and Hall (2000) referred that interdisciplinary exists between tourism and immigration. The literature tends to make the following points of convergence: i) tourism and the labour market in the host country, ii) tourism and migration, iii) tourism and entrepreneurship.

Clarke (2004) referred that there is a convergence between immigration and leisure activities in the host country. In turn, the link between tourism and immigration involves family and friends (Jackson, 1990, 2003 and Yuan et al. 1995). More recently, Lew and Wong (2002) demonstrated the importance of the internationalization of migration associated with the labour market (opportunities) and other forms of migration, the networks (family and friends, VFR). Tourism can be explained by international migratory movements. Tourism as a form of temporary international migration can be explained by the movements and structural changes in the distribution of the population.

Oigenblick and Kirschenbaum (2002) admitted that tourism is a facilitator of immigration.
In recent decades, the phenomenon of migration, trade and tourism has gained many adherents in academia as Kulendran and Wilson (2000), Eilat and Einav (2004), Phakdisoth and Kim (2007), Mervar and Payne (2007), Vogt (2008), and Fischer (2007). These studies showed that immigration and international trade seem to promote tourism.

The literature review of empirical tourism demand (VFR) (Crouch, 1994; Witt and Witt, 1995; Mervar and Payne, 2007; Lim, 1997 and Carrey, 1991) suggested that demand for tourism (dependent variable) can be measured by tourist arrivals. The number of overnight stays in hotel establishments by country of residence has also been used as the dependent variable.

In terms of explanatory variables, empirical models of tourism demand using income from tourists, service prices, exchange rates and geographical distance. The classic empirical studies such as Jud and Joseph (1974), Fuji and Mark (1981) and Carrey (1991) pointed out that it is the most important variable to explain the demand for tourism.

The link between immigration and tourism is usually explained by two trends: the push-pull forces and social capital. The decision is based mostly in the economic characteristics of the host country (exchange rates, market knowledge and prices). Economic factors affect the motivation of immigrants.

Foreign residents in a host country reduce transaction costs and develop networks. From the study of Hollander (1982) applied to the Australian case,
we can infer that the variable has a positive impact on migration with statistical significance.

A substantial attention has been given to the relationship between migrations, which has a significant influence on tourism arrivals. The questions from here to: Is international migrations sustaining VFR?

Other exogenous variables are also considered such as geographical distance or transportation costs (Phakdisoth and Kim, 2007; Allen and Yap, 2009, and Leitão 2010), infrastructures (Seetanah, 2006, and (Phakdisoth and Kim, 2007), population (Hanafiah and Haruin, 2010; Leitão, 2010), common language (Eilat and Einav, 2004).


Ledesma-Rodriguez et al. (2001) applied the panel data to analyze the short and long-run elasticises for visitors of Tenerife. The study of Naude and Saayman (2005) used a panel data for the period 1996-2000. The authors identified the determinants of tourism arrivals (VFR) in 43 African countries. Roget and Gonzalez (2006) studied the determinants for rural tourism demand in Galicia.
Maloney and Rojas (2005) used a dynamic panel data to analyse Caribbean destinations.


Recently, Mervar and Payne, (2007) used dynamic estimates to explain the determinants of tourism demand i.e. the lagged dependent variable. Also Phadisoth and Kim (2007) specify a panel of static data and dynamic (GMM-DIF) applied to the demand for tourism in Laos. They concluded that transportation costs, risk associated with the destination, bilateral trade and geographical distance are the main determinants of tourism in Laos.

Brida and Risso (2009) used a dynamic panel data study of the Germany demand for tourism in South Tyrol. The dynamic panel approach analyses the short and long-run effects. Brida and Risso (2009) concluded that the cost of travel and the relative price have a negative and significant impact on tourism demand. The authors also showed that the lagged dependent variable (tourism demand, VFR) has a positive and relevant effect on actual demand, reflecting according to the authors the loyalty of Germany tourists.

The study of Leitão (2010) specifies a static and dynamic panel data (GMM-System) for tourism demand in Portugal for the period 1995-2006. The GMM-System proposed by Blundell and Bon (1998, 2000) is an alternative to standard first differenced GMM estimator (GMM-DIF).
The author concluded bilateral trade, immigration, border, and geographical distance between Portugal and countries of origin are the main determinants of tourism demand.

4. Methodological approach and model specification

We use a regression to analyze the relationship VFR-migration. Before presenting the results of our estimations, we discuss the dependent variable and explanatory variables, describe the data model and address the hypothesis.

A gravity model will be used in estimating the relationship between immigration and international tourism to and from Portugal. Tourism is significant source of exports revenues for any country. Since the pioneering studies (Tinbergen, 1962; Poyhonen, 1963) pointed out that geographical distance is an important determinant of international trade. In this study we use a panel data methodology. The panel data analysis permits us to use more informative data and it accounts for unmeasured time-invariant determinants.

This study uses a dynamic panel data (GMM-System). In static panel data models, Pooled OLS, fixed effects (FE) and random effects (RE) estimators have some problems like serial correlation, heteroskedasticity and endogeneity of some explanatory variables.

The estimator GMM- system (GMM-SYS) permits the researchers to solve the problems of serial correlation, heteroskedasticity and endogeneity for some explanatory variables. These econometric problems were resolved by Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998,
who developed the first differenced GMM (GMM-DIF) estimator and the GMM system (GMM-SYS) estimator. The GMM-SYS estimator is a system containing both first differenced and levels equations. The GMM-SYS estimator is an alternative to the standard first differenced GMM estimator. To estimate the dynamic model, we applied the methodology of Blundell and Bond (1998, 2000), and Windmeijer (2005) to small sample correction to correct the standard errors of Blundell and Bond (1998, 2000). The GMM system estimator is consistent if there is no second-order serial correlation in the residuals (m2 statistics). The dynamic panel data model is valid if the estimator is consistent and the instruments are valid.

4.1 Econometric Model: Explanatory Variables and Data

Model Description

For the researchers, the visiting friends and relatives (VFR) are the most common variable used in creating econometric models of VFR, beside the total duration of visiting friends and relatives thousands of night per year\textsuperscript{1}. In this study, the regression of VFR in Portugal is from 16 different countries\textsuperscript{2} between the years 1995-2008. The data used to create the total number of visits, as dependent variable, are annually collected from INE- National Institute of Statistics. Our panel data is unbalanced.

\textsuperscript{1} In appendix we present the correlation matrix between variables.
\textsuperscript{2} The countries selected are Austria, Belgium, Brazil, Canada, Czech Republic, Denmark, France, Finland, Hungary, Italy, Luxembourg, Spain, Sweden, Netherlands, United Kingdom and USA.
First all the descriptive for panel data is presented in the following table.

**Table 1 Descriptive statistics for panel**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>LogVFR</td>
<td>5.80</td>
<td>0.57</td>
<td>4.51</td>
<td>6.89</td>
</tr>
<tr>
<td>LogVFR_{t-1}</td>
<td>5.77</td>
<td>0.59</td>
<td>4.52</td>
<td>6.89</td>
</tr>
<tr>
<td>LogGDP</td>
<td>4.28</td>
<td>0.03</td>
<td>4.23</td>
<td>4.36</td>
</tr>
<tr>
<td>LogIMI</td>
<td>3.16</td>
<td>0.68</td>
<td>1.57</td>
<td>5.03</td>
</tr>
<tr>
<td>LogPOP</td>
<td>7.28</td>
<td>0.26</td>
<td>5.61</td>
<td>8.46</td>
</tr>
<tr>
<td>LogDIST</td>
<td>3.52</td>
<td>0.78</td>
<td>2.80</td>
<td>3.86</td>
</tr>
</tbody>
</table>

Source: INE, Border Services, World Bank and authors’ calculation.

Following the literature review, we consider that visiting friends and relatives (demand for travel exports) in Portugal is a function of income, migration stock, population, and geographical distance.

\[
VFR_{it} = f(VFR_{it-1}, GDP, IMI, POP, DIST)
\]  

(1)

Where, VFR_{it} is the number of foreign tourist arrivals; VFR_{it-1} lagged dependent variable permits to analyse long run effects; GDP is the income in tourist generating countries; IMI is the number of migrants from various countries living in Portugal; POP is the total population in countries;

A series of hypothesis were formulated, considering how the selected variables will influence the effect of immigrant’s links to VFR tourism.
Hypothesis 1: There is a positive impact on the visiting friends and relatives in the long run.

Phakdisoth and Kim (2007), Brida and Risso (2009), and Leitão (2010) defended the idea that lagged tourism demand (VFR\(_{t-1}\)) has a positive impact on the economy.

Hypothesis 2: Tourism demand will be influenced by income of the tourist from countries of origin.

GDP, is the gross domestic product per capita in the country of origin of tourist, expressed in constant 2000 US$ was collected from the World Bank. According to the literature, we expect that the number of foreign tourist arrivals to increase in Portugal as the income in tourists’ origin country increase.

Hypothesis 3: Immigration flows play an important role in sustaining tourism.

The stock of immigration collected from the Border Services "Serviço de Fronteiras", (Ministry of Internal Affairs), corresponds to legal immigrants in Portugal.

According to previous studies, Dwyer et al. 1993, Seetaram 2008, and Oigenblick and Kirschenbaum (2002), Fischer (2007) tourism and immigration are correlated i.e. tourism encourages migration.
Hypothesis 4: Population changes in a country could positively sustain tourism flows

According to the literature (Witt and Witt, 1995; Oigenlick and Kirschenbaum, 2002; Hanafiah and Harun 2010) we expect a positive sign. The population data has been collected from World Bank.

Hypothesis 5: Tourism increases if transportation cost decreases.

DIST is the geographical distance between Portugal and the tourist generating countries. According to the literature we expect a negative sign for this variable.

4.2 Model Specification

\[ VFR_i = \beta_0 + \beta_1 X_i + \xi_i + \eta_i + \varepsilon_i \]  

Where \( VFR_i \) is the total number visits by foreign nationality to Portugal, \( X \) is a set of explanatory variables. All variables are in the logarithm form; \( \eta_i \) is the unobserved time-invariant specific effects; \( \xi_i \) captures a common deterministic trend; \( \varepsilon_i \) is a random disturbance assumed to be normal, and identical distributed (IID) with \( \text{E}(\varepsilon_i) = 0; \text{Var}(\varepsilon_i) = \sigma^2 > 0 \).

The model can be rewritten in the following dynamic representation:

\[ VFR_i = VFR_{i-1} + \beta_0 + \beta_1 X_i - \rho \beta_1 X_{i-1} + \xi_i + \eta_i + \varepsilon_i \]  

(10)
4.2.1 Times-Series Properties of the Variables in the Panel

Before estimating the panel regression model, we realize a test for unit root of the variable. Following the literature we apply a battery of unit root tests: ADF-Chi square, and PP-Chi square. The table 2 presents the results of different panel unit root test (ADF-Fisher Chi square, PP- Fischer Chi square).

<table>
<thead>
<tr>
<th>Panel unit root test of variables</th>
<th>intercept and trend</th>
<th>Statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>LogVFR</td>
<td></td>
<td>ADF-Fisher Chi-square</td>
<td>63.7344</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP-Fisher Chi-square</td>
<td>87.5523</td>
</tr>
<tr>
<td>LogGDP</td>
<td></td>
<td>ADF-Fisher Chi-square</td>
<td>53.1586</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP-Fisher Chi-square</td>
<td>147.6657</td>
</tr>
<tr>
<td>LogIMI</td>
<td></td>
<td>ADF-Fisher Chi-square</td>
<td>374.2890</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP-Fisher Chi-square</td>
<td>118.896</td>
</tr>
<tr>
<td>LogPOP</td>
<td></td>
<td>ADF-Fisher Chi-square</td>
<td>108.8491</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP-Fisher Chi-square</td>
<td>96.3302</td>
</tr>
<tr>
<td>LogDIST</td>
<td></td>
<td>ADF-Fisher Chi-square</td>
<td>16.0486</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP-Fisher Chi-square</td>
<td>36.3003</td>
</tr>
</tbody>
</table>

Source: INE, Border Services, World Bank and authors’ calculation.
The most important model variable such as the visit friend (LogVIST), income per capita (LogGDP), stock of immigration (LogIMI), and population (LogPOP) do not have unit roots, i.e are stationary, with individual effects and individual trend specifications.

5. Empirical Results

This section presents the estimation using GMM-System estimator proposed by Arellano and Bover, (1995) and Blundell and Bond, (1998, 2000). We used STATA econometric software to estimate the model. The model presents consistent estimates, with no serial correlation the Arellano and Bond test for $M_2$. The specification Sargan test shows that there are no problems with the validity of instruments used. The GMM system estimator is consistent if there is no second-order serial correlation in the residuals ($M_2$ statistics). The dynamic panel data are valid. The Windmeijer (2005) finite sample correction is used. In the table 3 we can observe the relationship between tourism demand and immigration. The general performance of the equations is satisfactory. All explanatory variables are significant (LogGDP, LogIMI, LogPOP, and LogDIST at 1% level); the coefficient (LogVFR_{t-1}) is significant at 5% level. For Lagged dependent variable (LogVFR_{t-1}), the expected sign is positive confirmed by the results. So we can infer that this variable has a positive impact.
on Portuguese economy. Brida and Risso (2009), and Leitão (2010) also found a positive sign for lagged dependent variable.

As expected, the variable LogGDP has a significant and positive effect on tourism demand. Phakdisoth and Kim (2007), and Brisa and Risso (2009) also found this result.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>Expected Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>LogVFR(_{t-1})</td>
<td>0.101 (2.26)**</td>
<td>(+)</td>
</tr>
<tr>
<td>LogGDP</td>
<td>1.41 (5.67)***</td>
<td>(+)</td>
</tr>
<tr>
<td>LogIMI</td>
<td>0.49 (9.27)***</td>
<td>(+)</td>
</tr>
<tr>
<td>LogPOP</td>
<td>0.15 (5.05)***</td>
<td>(+)</td>
</tr>
<tr>
<td>LogDIST</td>
<td>-1.05 (-5.98)***</td>
<td>(-)</td>
</tr>
<tr>
<td>C</td>
<td>13.50 (9.25)***</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>-1.43 [0.153]</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>0.97 [0.32]</td>
<td></td>
</tr>
<tr>
<td>Sargan test</td>
<td>15.91 [1.00]</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>125</td>
<td></td>
</tr>
</tbody>
</table>

The null hypothesis that each coefficient is equal to zero is tested using one-step robust standard error. T-statistics (heteroskedasticity corrected) are in round brackets. P-values are in square brackets; ***/**/*- statistically significant at the 1 per cent, 5 per cent, and 10 per cent levels. AR (2) is tests for first-order and second–order serial correlation in the first-differenced residuals,
asymptotically distributed as N(0,1) under the null hypothesis of no serial correlation (based on the efficient two-step GMM estimator). The Sargan test addresses the over-identifying restrictions, asymptotically distributed $X^2$ under the null of the instruments’ validity (with the two-step estimator).

Source: INE, Border Services, World Bank and authors’ calculation.

The variable the migrant stock (LogIMI) is positively related to the dependent variable. Oigenblick and Kirschenbaum, (2002); Fischer, (2007) and Seetaram, (2008) showed that the level of tourism depends not only on the population of origin country, but also on the migration flows.

The variable population (LogPOP) finds a positive sign, as we expected, and corresponds to the results of Hanafiah and Harun (2010), and Leitão (2010). A 1% increase in population of the origin country would increase 0.15% foreign tourist arrivals to Portugal.

The coefficient of LogDIST (Distance) validates the hypothesis 5. This result confirms the importance of the neighbourhood. Following Phakdisoth and Kim (2007), we can conclude that international tourism demand is directly influenced by the distance from the countries of origin of tourists and tourism destination country.

6. Conclusions

In this paper we analyzed the relationship between VFR visits and
immigration using dynamic panel data in case of Portugal. The GMM –system has rarely been applied to tourism analysis. Our results supported the hypothesis that there is a positive correlation between immigration and tourism demand. This result is in line with in existing literature such Dwyer et al. (1993); Oigenblick and Kirschenbaum, (2002), Fischer (2007) and Seetaram, (2008). The econometric models showed that GDP per capita, and population which determines the ability to travel, these explanatory variables have a positive impact on VFR visit. For geographical distance, the results indicated that it has a negative influence on inflows, as expected. Generally, the econometric model is in line with the results with previous empirical studies. Finally, although the use of more recent econometrical techniques should at least be compared, and it would be dangerous to generalize from this empirical study, it may be preferable to use the GMM approach in empirical visiting friends and relatives (VFR) or tourism demand, rather than static panel data (pooled OLS, fixed effects, and random effects).

Appendix

Table 4 Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>LogVFR</th>
<th>LogGDP</th>
<th>LogIMI</th>
<th>LogPOP</th>
<th>LogDIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>LogVFR</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LogGDP</td>
<td>-0.06</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LogIMI</td>
<td>0.57</td>
<td>0.03</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LogPOP</td>
<td>0.13</td>
<td>0.16</td>
<td>0.30</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>LogDIST</td>
<td>0.06</td>
<td>0.14</td>
<td>0.16</td>
<td>0.20</td>
<td>1.00</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>130</td>
</tr>
</tbody>
</table>
Source: INE, Border Services, World Bank and authors’ calculation.

**Figure 2** Distribution of VFR

Source: Portuguese Border Services and authors’ calculation.
Kernel density estimate

Source: Portuguese Border Services and authors’ calculation

Figure 3 Dependent variable Kernel density

References


