Significance bias in the tourism-led growth literature

Nino Fonseca
Polytechnic Institute of Viana do Castelo, Portugal

Marcelino Sánchez-Rivero
University of Extremadura, Spain

Abstract
We use an original meta-regression analysis to test the existence of bias of statistical significance in the literature on the study of Granger causality relationships between tourism and income. We conclude for the presence of such bias. We also conclude that some methodological choices are more likely to lead to statistically significant results. Additionally, we find that the empirical association between tourism and income is stronger the shorter the time horizon analyzed and that it is correlated with the countries’ economic, touristic and demographic profile. Altogether, our results suggest that there are reasons to mistrust the typical methodological approach followed by the literature reviewed and cast doubt on its usual implications for economic policy.

Keywords
Granger causality, meta-regression, publication bias, tourism-led growth

Introduction
The pressure to publish or perish, ubiquitous in the academia, coupled with the seeming preference of scientific journals’ editors for the publication of studies with statistically significant results, has motivated extensive discussions about the reliability of the results of the scientific enterprise in all areas of knowledge (Carafoli, 2015; Ioannidis, 2012; Lacetere and Zirulia, 2011; Necker, 2014). As such, it is appropriate to ask whether the results of the line of research initiated by Balaguer and Cantavella-Jordá (2002) represent a genuine empirical effect between tourism and income.

We are concerned with the analysis of the literature on the study of the tourism-led growth hypothesis based on Granger causality tests. Three recent and quite complete narrative reviews

Corresponding author:
Nino Fonseca, Superior School of Technology and Management, Polytechnic Institute of Viana do Castelo, Avenida do Atlântico, s/n, 4900-348 Viana do Castelo, Portugal.
Email: ninomf@estg.ipvc.pt
have already been presented by Brida et al. (2013), Pablo-Romero and Molina (2013), and Brida et al. (2016). Our aim here is to evaluate the presence of a specific type of publication bias, the so-called bias due to statistical significance (Ashenfelter et al., 1999; Görg and Strobl, 2001). This type of bias occurs when there is an overrepresentation of statistically significant empirical effects. In these circumstances, the set of published studies tends to reveal an empirical effect that is apparently, albeit deceptively, higher (in absolute value).

Doucouliagos et al. (2005) suggest that in areas where economic theory is consensual (e.g. negative price elasticity, positive returns from education, etc.), it is more likely to find problems of publication bias. Where it is possible to find theoretical support for both positive and negative effects, issues of publication bias are less likely to be found because any empirical results are always susceptible of theoretical interpretation (e.g. the empirical literature on the study on the role of democracy in economic growth). However, those areas of study may be contaminated by problems of bias due to statistical significance precisely because both empirical effects, positive or negative, are amenable to economic interpretation. This is the case of the empirical analysis of Granger causality relationships between tourism and income because it is possible to interpret the evidence found either in one direction (from tourism to income) or the other (from income to tourism).

Is the bias due to statistical significance a reason for concern in the empirical literature initiated by Balaguer and Cantavellá-Jorda (2002)? This is the question we seek to answer in this article. The presence of this kind of bias means that authors of articles and journal editors tend to prefer statistically significant results, with negative consequences for both the accrued scientific knowledge and for the consolidation of scientific practices, as well as for the legitimacy of the recommendations of economic policy that usually stem from those results.

To answer this question, we carry out a meta-analysis of an empirical effect that represents the degree of association between tourism and income and is independent of the causal directions possibly detected through the Granger causality tests. At the same time, we evaluate the genuineness of this empirical effect. We also test the extent to which variables presumably correlated with publication bias and variables likely to explain the variability of the association between tourism and income add to the explanation of the heterogeneity of the empirical effect. We pursue two objectives here. The first is to identify the sources of the bias due to statistical significance. The second is to understand why the association between tourism and income is stronger in some cases than in others, regardless of the dominant Granger causal direction.

Our methodological basis is the meta-regression analysis proposed by Stanley and his coauthors (Stanley and Doucouliagos, 2012; Stanley et al., 2013). Given the specificities of the Granger causality tests, we make several modifications aimed to deal with the nature of the corresponding empirical effects. In this sense, we present an innovative methodological contribution that is relevant both to our meta-analysis and to the broader scope of empirical research in social sciences. Notably, Bruns et al. (2014) inaugurate the application of methods of meta-regression analysis to the results of Granger causality tests, suggesting strategies to test for the presence of publication bias and the genuineness of the empirical effects. We contribute to this line of research by proposing a meta-regression equation intended to test for the presence of bias due to statistical significance when the researcher is concerned with the results of Granger causality tests.

The remaining of this article consists of four additional sections. In the following section, we describe the methodology that will be applied to answer the questions raised above, together with the empirical model and the estimation procedure. The data set is described in the third section, and
the results of the meta-regression analyses are carried out along the fourth section. The fifth and last section presents our main conclusions.

Methodology of the meta-regression analysis

Modifications of the standard methodology

The bias due to statistical significance occurs when the authors of studies with small samples manipulate their econometric specifications to find larger estimates for the effect size. Since small samples tend to be associated with higher standard deviations, this type of publication bias implies a positive relationship between the estimates of the effect size of each empirical study and their respective standard deviations (Ashenfelter et al., 1999; Görg and Strobl, 2001). The intuitive explanation is as follows: If the sample size grows as the standard deviation increases, it could only happen if the sample or the econometric specification is being manipulated to find a large effect size. Why? Because we know, from statistical theory, that larger standard deviations are associated with smaller effect sizes, not larger ones.

In the regression aimed to test for the bias due to statistical significance, the dependent variable is the modulus or absolute value of the empirical effect, $\text{Effect}_i$, and the independent variable corresponds to the respective standard deviation, $SE_i$:

$$|\text{Effect}_i| = \beta_A + \alpha_A SE_i + \varepsilon_i.$$  \hspace{1cm} (1)

The term $\beta_A$ corresponds to the intercept and $\varepsilon_i$ is a random disturbance term. The null hypothesis analyzed is that of the absence of bias ($\alpha_A = 0$) and the alternative hypothesis is the presence of bias ($\alpha_A > 0$). The rejection of the null hypothesis means that the empirical effects presented in the literature tend to be overestimated due to the preference of authors and editors for statistically significant results. Nevertheless, the possibility of heteroscedasticity requires, instead, the estimation of the meta-regression given below, where the meaning of the tested hypotheses remains:

$$|t_i| = \alpha_A + \beta_A (1/SE_i) + \mu_i.$$  \hspace{1cm} (2)

Equation (2) results from the correction of heteroscedasticity through the division of both sides of equation (1) by the standard deviation, $SE_i$. Therefore, the dependent variable in equation (2) is the absolute value of the well-known $t$ statistic. The intuitive interpretation is that as the standard deviation increases and the term $\beta_A (1/SE_i)$ vanishes, no effect size would remain either ($\alpha_A = 0$). If it remains ($\alpha_A > 0$), it must be due to the fact that the sample or the econometric specification is being manipulated to find such non-null effect sizes.

However, when the meta-analysis concerns the results of Granger causality tests, three modifications are necessary. Firstly, for the results of these tests to be used as empirical effects, it is required to convert them into a common metric, with a standard distribution and with properties suitable for a meta-regression analysis. Bruns et al. (2014), in line with Stanley (2005), and as suggested by Abramowitz and Stegun (1964), convert the probabilities of significance of the $F$ and $\chi^2$ test statistics into normal standardized variables through the symmetric of a probit function, which corresponds to the inverse of the standard normal distribution. This gives rise to a metric that associates the larger values with the likelier probabilities of rejection of the null hypothesis of Granger noncausality. For details, see Bruns et al. (2014).
the coefficient of the dummy variable *Tugcu (2014)*, which captures the fact that the variable that represents the real income is a growth rate, has a coefficient which is always negative, and, in most of the cases, statistically significant at a significance level of 5% or 10%. Looking at the results with the full sample (columns 1 to 4 of Table 6), without the dummy variable *Tugcu (2014)*, the dummy variable Real GDP comes out positive and statistically significant (columns 1 and 2 of Table 6); when the dummy variable *Tugcu (2014)* is included, it comes out negative (column 3) and even highly statistically significant after the process of reduction of variables (column 4; *p*-value < 0.0001), at the expense of the statistical significance of the dummy variable Real GDP. This means, quite clearly, that the studies that analyze levels of income instead of the respective growth rates are associated with higher empirical effects.

When it comes to the variable *Johansen*, it is not present in the preferred meta-regression, even though its coefficient turned out to be consistently positive and, on some occasions, statistically significant at significance level of 5% or 10%, in the additional meta-regressions we performed. In turn, the coefficients of the variables that represent the options regarding the measurement of tourism and the type of Granger causality test are statistically significant twice (columns 1 and 4) but only for significance levels of 10%.

**Economic, touristic, and demographic profile**

The results of our preferred meta-regression (column 4 of Table 6) show us that all the coefficients of the continuous independent variables intended to capture the variability of the empirical effect—namely the level of *Economic development*, the degree of *Tourism specialization*, and the *Country size*—are statistically significant for levels of significance lower than 5% (respectively, *p*-value = 0.0003, *p*-value < 0.0001, and *p*-value = 0.0863). The conclusion is qualitatively similar in columns 1 to 3. These variables only turn out to be statistically irrelevant in columns 5 and 6, albeit without changes in the signs of the coefficients where they emerge. In the additional meta-regressions not presented here, the signs obtained for these variables were, once again, systematically positive and, more than 75% of the times, statistically significant for levels of significance below 5% or 10%.

**Conclusion**

In this article, we tested for the presence of bias due to statistical significance in the empirical literature concerned with the study of Granger causality between tourism and income, as suggested initially by Balaguer and Cantavella-Jordá (2002). For this purpose, we based our work on an original approach where the empirical effect meta-analyzed consisted of an indicator of the degree of association between tourism and income, built from the results of Granger causality tests but independent of the causal directions detected. The approach developed is a contribution that goes beyond the scope of this study as it is applicable in any other meta-regression studies intended to test for the presence of bias due to statistical significance based on the results of Granger causality tests. We found evidence for the presence of such bias in the literature of our concern. Nevertheless, because the statistical properties of our test, namely regarding size and power, are not studied yet, this approach can only be taken as a description of the data and of the relationships between the variables.

At the same time, we sought to find out whether the empirical effect or empirical association is genuine or not, that is, whether it reflects the existence of a statistically consistent relationship between tourism and income. We concluded that the empirical effect exists, but it is not genuine,
because the Granger causal association between tourism and income weakens over time. In other words, the value of the empirical effect tends to decrease as the sample dimension—that is, the time span—increases. This fact is in odds with what we would expect if the empirical effects were statistically consistent (and genuine). This conclusion may mean that there are problems of publication bias associated with the deliberate manipulation of sample sizes to obtain statistically significant results, or it may mean, additionally or not, that the Granger causal effect between tourism and income is confined to short time horizons. Thus, the analysis of shorter samples increases the chances of obtaining larger and more statistically significant empirical effects.

Finally, we evaluated the extent to which the variables correlated with the publication bias and the variables that represent the economic and demographic characteristics of the countries included in our sample contribute to the explanation of the variability of the empirical effect.

Our findings show that the measurement of income through the level of real GDP adds significantly to the achievement of higher empirical effects. On the other hand, the choice of a growth rate has the opposite implications. There is also some evidence that the options regarding cointegration analysis are slightly relevant to the explanation of the variability of empirical effects. This means that the options related to the research design are likely to contribute to obtaining larger or smaller empirical effects.

We concluded, as well, that the richer, more tourism-specialized and more populous countries are also those with a stronger degree of Granger causal associations between tourism and income. However, in the absence of a much more robust theoretical support and confirmation through further statistical and econometric analyzes that mitigate the shortcomings of the standard approach, it is not advisable to move forward with concrete, practical implications.

Obviously, we do not intend our study to be definite about the questions raised. Being a quantitative review of the literature, it should be replicated as soon as new studies come to light. This means that our conclusions are confined to the available sample of studies. Nevertheless, taken together, these conclusions suggest that there are reasons to be somewhat skeptical about the methodological approach of the literature reviewed and cast doubts on their findings and implications for economic policy. Naturally, the existence of a collective significance bias does not deny the value of individual studies where significant effects are properly detected. These studies do offer useful policy or managerial implications for the specific countries analyzed.

The next step will be to confirm our findings in the context of related lines of research and alternative data structures. This is so because, on the one hand, there is a wide range of studies that also analyze the long-term relationships between tourism and income based on time series but does not apply Granger causality tests. On the other hand, there is a wide set of empirical studies that falls within the tradition of growth econometrics with cross-sectional data and panel data, where tourism emerges as a determinant of economic growth, which needs to be reviewed using meta-regression analyses. If our conclusions find confirmation, it will be necessary to justify theoretically the empirical regularities found and to rethink the empirical methodologies aimed at analyzing the role of tourism in economic growth.

**Declaration of conflicting interests**
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**
The author(s) received no financial support for the research, authorship, and/or publication of this article.