Population and tourism drivers of Spanish fishing communities (FC)

Ikerne del Valle

Detp. Public Policies and Economic History.

University of the Basque Country.

Extended Abstract

Fishing communities (FC) are singular. Not only are they located in privileged coastal natural environments, but they also offer long-standing cultural traditions and attractive cultural heritage frameworks, that, overall, make them a driving target for population and tourism flows. To empirically explore the latter singularity, we will focus on the subsample of Spanish coastal communities. We are excluding non-coastal municipalities from the analysis to avoid the bias that the closeness to the sea would for sure generate. The dummy variable coastal community typology (CCT) takes the value 1 if the municipality is a FC, and zero if it is a non-fishing coastal community (nFC). Moreover, taking advantage of del Valle and Astorkiza (2021) we are using the a score of Spanish FCs using composite indicators (CI) calculated from a mixture of desirable (DI) and undesirable (UI) individual indicators making use of the robust extension of the Benefit of the Doubt (BoD) approach, a nonparametric method based on data envelopment analysis (DEA) that, in the context of productive efficiency, provides a data oriented endogenous weighting method that can be regarded as an input-oriented constant returns to scale CRS-DEA model, with all individual indicators considered as outputs, and a single input equal to one for all inputs. The outcome of such a procedure is a CI that provides a score to infer the ranking of the FCs. DI include the value of the landings (LAN), the number of vessels (NV), the gross tonnage (GT), the number of fishers (NF) and the number of vessels under 10 years (NEW), while the average age of the fleet (AGE) is an UI.

Both FC and nFC share the valuable link to the sea, but only around 29% of the total Spanish coastal communities may be catalogued as FCs according to the operative definition adopted in section 2. Our hypothesis is that, FC differ from nFC, although, simultaneously, they may exhibit an heterogenous performance and idiosyncrasy depending on their geographical location (GEO) and proximity (or not) to densely populated areas (FUA). In addition to the main effects of CCT, GEO and FUA, the three factors may well interact to exert additional joint effects on key response variables related to demography (VPOP, AGING, DEPENDENCE) and tourism (PTA, TA, TOUR).

The aim of this paper is to explore the interplay among the above mentioned three independent factors, namely, coastal community typology (CCT), geographical location (GEO) and the fact of being within the commuting zone of a densely city (FUA), so as to assess whether they interact to explain key individual dependent variables related to demography (VPOP, AGING, DEPENDENCE) and tourism (PTA, TA, TOUR). Since

our primary concern is the effect of coastal fishing community typology, CCT will be the focal variable in the analysis. Nevertheless, we presume that any of the above-mentioned dependent variables may also be influenced by GEO and FUA (the two moderator variables) and by the interaction among the three factors (CCT:GEO:FUA). Methodologically, the basic one-way ANOVA that splits the target population in FC and nFC, will be followed by a three-way ANOVA design so as to verify whether there is a statistically significant difference between the means of different groups that have been split on CCT, GEO, and FUA; and to determine how such factors interact and affect demographic and touristic variables.

The underlying three-way ANOVA procedure applied for each of the demographic and touristic variables in the set X={VPOP, AGING, DEPENDENCE, PTA, TA, TOUR} may be summarized as follows. Starting from the model that includes main and interaction effects (CCT*GEO*FUA), if there is a significant three-way interaction effect (CCT:GEO:FUA), we will decompose it into simple two-way interactions by running two-way interaction at each level of the third factor. Hence, we will evaluate: a) the effect of GEO*FUA interaction at each level of CCT; b) the effect of CCT*GEO interaction at each level of FUA; c) the effect of FUA*CCT interaction at each level of GEO. Equivalently, if the three-way interaction is rejected, we will focus on determining whether there is any statistically significant two-way interaction, follow up by simple main effects analyses if two-way interactions is rejected. Additionally, we will also study pairwise comparisons between selected subgroups.

Since the nature of our data is unbalanced (i.e. the number of coastal municipalities in each subgroup is different) attention should be paid on the method to split the total variation of any of the dependent variables in X (measured as sums of squares) into different sources of variation. There are three fundamentally different approches to run an ANOVA in an unbalanced design, commonly called Type I, II and III sums of squares. Contrary to what happens with balanced data, when working with unbalanced data inferential statistics from using Type I, Type II or Type III sums of squares are not equal, which raises the ongoing controversial question of the error type choice. See among others, Her (1986), Stewart-Oaten (1995) and Smith and Cribbie (2014) for a detailed discussion on the topic. ANOVA based on Type I sum of squares follows a sequential pattern and will give different results depending on the factor order, which is undesirable[1]. Both ANOVA based on type II and III sum of squares are not sequential, so the order of specification of the factors does not matter, but they differ on their underlying assumptions about the interactions among the factors and the principle of marginality. Type-II sums of squares are constructed obeying the principle of marginality, so the hypotheses tested assume that terms to which a particular term is marginal are zero and do not take an interaction effect. Type-III tests do not assume that terms higher-order to the term in question are zero and, thus violates marginality; and unlike Type II, the Type III sums of squares do specify an interaction effect. Generally, Type II sum of squares has greater statistical power, and therefore is the most appropriate default if there is no evidence of an interaction (Langsrud, 2003). However, if a large interaction exists in the population Type III method may hold better statistical power, but the main effects will be of dubious value (anyway, in the presence of interactions, main effects are rarely interpretable). Since we expect the interaction among the factors to be significant, we decided to use Type III error, by means of contrasts that are orthogonal in the row-basis of the model matrix (In R, we are using contr.sum). Because the multi-way ANOVA

[1] This approach in fact tests for a difference in the weighted marginal means, which in practical terms means that the results are dependent on the proportions in the particular data set.

model is over-parametrized, it is necessary to choose a contrast setting that sums to zero, otherwise the ANOVA analysis will give incorrect results with respect to the expected hypothesis.

The results show that FC and nFC differ in five out of six indicators in X. Based on one-way ANOVA, there is a significant statistical difference in the average PTA, TA, TOUR, VPOP and AGE values of FC and nFC. The difference in DEP is not statistically significant. {PTA: [F(1,801)=106 (0.0000), ges=0.117]; TA: [F(1,801)=111.2 (0.0000), ges=0.122]; TOUR: [F(1,801)=87.8 (0.0000), ges=0.10], VPOP: [F(1,797)=25.98 (0.0000), ges=0.03]; AGE: [F(1,801)=4.13 (0.042), ges=0.003]; DEP: F(1,801)=2.43 (0.12); ges=0.003]}. The related mean, standard error (se) and 95% confidence interval (CI) values for FC and nFC are respectively {PTA(FC): 3350, se=134, CI=[512,1038]; PTA(nFC): 775, se=211, CI=[2936, 3764]; TA(FC): 680, se=43, CI=[595,764]; TA(nFC): 143, se=27, CI=[89,196]; TOUR(FC): 46915, se=3249, CI=[40537, 53293]; TOUR(nFC): 10850, se=2065, CI=[6797,15903]; VPOP(FC): 6.98, se=1.457, CI=[4.12,9.84]; VPOP(nFC): 15.78, se=0.926, CI=[13.96,17.60]; AGE(FC): 160, se=5.66, CI=[148, 171]; AGE(nFC): 146, se=3.6, CI=[139, 153]; DEP(FC): 55.7, se=0.38, CI=[54.9, 56.4]; DEP(nFC): 54.5, se=0.598, CI=[53.4, 55.7]}. Summarizing, average values of touristic indicators are higher for FC than for nFC, while demographic ones are below. However, these results may be misleading because we are ignoring the potential effects of FUA and GEO and the potential interaction among the factors.

In fact, the interaction plots suggest that, not only GEO but also FUA might largely influence on the level of the demographic and touristic indicators {PTA, TA, TOUR, VPOP, AGE, DEP}. For instance, the positive gap in PTA between FC and nFC is more pronounced in the E, S and IC compared to NO and NE, and, in general FUA1 municipalities exhibit greater PTA levels than FUA0 ones. The difference in TA between FC and nFC is also more pronounced in the E, S and IC compared to NO and NE. Besides, in FCs for all GEO levels except IC, FUA1 municipalities exhibit greater TA levels than FUA0 ones. The gap between FUA1 and FUA0 is positive for all GEO levels except IC, and the difference noticeably pronounced for FCs. TOUR follows similar profiles, with a gap between FC and nFC substantially more pronounced in the E and S compared to NO, NE and even IC and in FCs for all GEO levels except IC, FUA1 municipalities exhibit greater TOUR levels than FUA0 ones. The difference in TOUR between FUA1 and FUA0 is positive for all GEO levels except IC, and noticeably more pronounced for FC. The gap in VPOP between FC and nFC is more pronounced in S compared to E, NO and NE and IC is the only GEO location showing a negative gap; and on the other, in FC for all GEO levels except IC, FUA1 municipalities exhibit greater VPOP levels than FUA0 ones. The gap between FUA1 and FUA0 is positive for all GEO levels except IC and the difference is less pronounced for FC. Aging is more pronounced in NO, with twice the average of S (the location with the average minimum AGE). On the other hand, there is a positive gap between FUA0 and FUA1 (more pronounced in the NO). Contrary to E, IC and S, FCs in NE and NO exhibit somewhat higher levels of AGE than nFC ones. Dependence is more pronounced in NO, NE and E (more than ten points in average than S and IC (the location with the minimum average DEP). On the other hand, there is a positive gap between FUA0 and FUA1 (more pronounced in NO, NE and E). In general FCs show lower average dependence levels. Contrary to E, IC and NE, NO and S exhibit somewhat higher levels of DEP than nFC ones.

There was a statistically significant simple two-way interaction between CCT and GEO (CCT:GEO) for FUA0, $F(4, 783) = 2.79$, $p = 0.026$, but not for FUA1, $F(4, 783) = 0.97$, $p = 0.41$. This result suggests that for FUA0 the effect of CCT on AGING, depends on

GEO. b) There was a statistically significant simple two-way interaction between GEO and FUA (GEO:FUA) for FC, $F(4, 783) = 3.33$, $p = 0.019$, but not for nFCC, $F(4, 783) = 2.35$, $p = 0.053$. This result suggests that for FCs the effect of FUA on AGING depends on GEO. c) There was a statistically significant simple two-way interaction between CCT and FUA (CCT:FUA) for E, $F(1, 783) = 5.54$, $p = 0.01$, and S $F(1, 783) = 5.52$, $p = 0.019$ but not for IC $F(1, 783) = 5.54$, $p = 0.69$, NE $F(1,783)=0.026$ $p=0.87$ and NO $F(1,783)=0.012$ $p=0.912$. This result suggests that the effect of CCT on AGING depends on FUA only for E and S. Note that, statistical significance of a simple two-way interaction was accepted at a Bonferroni-adjusted alpha level of 0.025. This corresponds to the current level you declare statistical significance at (i.e., $p < 0.05$) divided by the number of simple two-way interaction you are computing (i.e., 2). A statistically significant simple two-way interaction can be followed up with simple simple main effects. In our example, you could therefore investigate the effect of treatment on pain score at every level of risk or investigate the effect of risk at every level of treatment. Group the data by GEO and FUA and analyze the simple simple main effects of CCT on each of the variable X={PTA, TA, TOUR, VPOP, AGING, DEPENDENCE}.

The two-way interaction CCT:GEO was statistically significant for both FUA0 and FUA1 municipalities for: PTA {FUA0 [F(4, 783)=8.23 (0.0000), ges=0.04)]; FUA1 [F(4, 783)=6.58 (0.0000), ges=0.033)]}; TA {FUA0 [F(4, 783)=6.08 (0.0000), ges=0.03)], FUA1 [F(4, 783)=7.31 (0.0000), ges=0.036)]}; TOUR { FUA0 [F(4, 783)=3.72 (0.0000), ges=0.019)], FUA1 [F(4, 783)=12.2 (0.0000), ges=0.059)]}; and VPOP {FUA0 [F(4, 779)=2.48 (0.043), ges=0.013)], FUA1 [F(4, 779)=4.55 (0.0045), ges=0.023)]}. The two-way interaction CCT:GEO was statistically significant for FUA0 municipalities but not significant for FUA1 ones for AGING {[F(4, 783)=2.79 (0.0000), ges=0.187)], [F(4, 783)=0.979 (0.418), ges=0.005)] and DEP {[F(4, 783)=4.01 (0.003), ges=0.02)], [F(4, 783)=0.362 (0.836), ges=0.002)].

The only two-way significant GEO:FUA interaction for FUA0 municipalities concerns VPOP [F(4, 779)=0.18 (0.418), ges=0.015)]; whereas for FUA1 municipalities a significant GEO:FUA interaction has been found for AGING [F(4, 783)=3.33 (0.001), ges=0.017)], PTA [F(4, 783)=3.91 (0.003), ges=0.02)], TA [F(4, 783)=3.80 (0.005), ges=0.019)] and TOUR [F(4, 783)=3.78 (0.005), ges=0.019)].

Significant two-way CCT:FUA interactions have been found in E {VPOP [F(4, 779)=4.25 (0.04), ges=0.005)], AGING [F(4, 783)=5.54 (0.046), ges=0.007)], DEP [F(4, 783)=7.58 (0.006), ges=0.01)], TA [F(4, 783)=4 (0.046), ges=0.005)], TOUR [F(4, 783)=19.7 (0.0000), ges=0.025)]} and S {VPOP [F(4, 783)=10.1 (0.002), ges=0.013)], AGING [F(4, 783)=5.52 (0.019), ges=0.007)]}.