Are municipalities in the red to go green? The Blue Flag case.

Fabio I. Martinenghi

1 Introduction

As Buckley (2002) puts it "ecolabels and environmental accreditation are controversial topics in tourism". Indeed, the problem is more complex than it seems. On the one side, there is no consensus around a precise definition of "ecolabel", which has become an umbrella term associated with a variety of certification types: mandatory, compulsory, multidimensional, monodimensional, etc. (Jha et al., 1997). On the other side, the effectiveness of ecolabels cannot be generalised, because it depends on the features of each single one, together with how informed the consumers are on the specific issue the label tackles and how strong its reputation is (Thøgersen et al., 2010). Given these considerations, it should not come as a surprise that the evidence around the impact of ecolabels, and more particularly beach awards (McKenna et al., 2011), is mixed. In light of these external validity issues, it is important to focus on the assessment of popular ecolabels since they affect large areas and promise to bring together financial and wildlife conservation interests , to the benefit of all stakeholders.

In this paper, I focus on one of the most famous and widely-adopted ecolabel, the Blue Flag programme. In particular, I look at how being awarded for the first time[1] affects a municipality's balance sheet and its supply of collective tourist accommodation. I focus on first-time certifications and their temporary effects because I expect the maximum impact of the Blue Flag to be reached when the certification is new. Indeed, a first-time certification is celebrated locally by the municipality, nationally by traditional media and internationally on the Blue Flag programme's website (www.blueflag.global). Moreover, the effect is likely to be temporary in countries like Spain, Greece, France and Italy where a large fraction of coastal municipalities have already been awarded in the past decades. In particular, due to this Blue Flag crowdedness, the effect is likely attributable to the extra-ordinary positive media exposure a municipality receives after a first-time certification.

I adopt a pooled event study approach using Italian data between 2002-2016 to evaluate the effects of a first-time Blue Flag certification on the awarded municipalities. I find that

[1]Throughout the paper, 'municipalities awarded for the first time' is intended as 'municipalities that are observed passing from not-certified to certified for the first time within the sample'. Lack of data on Blue Flag certifications prior to 2002 prevents me from being sure that any in-sample first certification is the first certification.

municipalities awarded with a Blue Flag for the first time significantly increase their revenues, while I find no evidence that the award leads to an increase in the supply of collective tourist accommodation supply. My findings also provide further evidence (see Creo and Fraboni, 2011; Pencarelli et al., 2016; Cerqua, 2017) that, as proposed by Zielinski and Botero (2015), the Blue Flag award is an opportunity for mayors to promote and enact environmentally-conscious infrastructural improvements.

The literature on environmental beach certifications in tourist destinations is either survey (McKenna et al., 2011), or focusing on tourism in developing countries (Blackman et al., 2014), or looking at tourists' flows (Blanco et al., 2009; Capacci et al., 2015; Cerqua, 2017), or a combination of these and it focuses on the Blue Flag programme. For an updated review on the Beach Certification Schemes and the Blue Flag literature, see Zielinski and Botero (2019).

The findings on the economic effects of the Blue Flag award are not unequivocally significant and positive. At a survey level, McKenna et al. (2011) finds that beachgoers in Ireland, Wales, Turkey and the USA do not choose to visit a beach based on beach awards. However, Blackman et al. (2014) reports that the Blue Flag certification led to "19 new hotels and 1628 new hotel rooms per year" in Costa Rica. In developed countries such as Italy (see Capacci et al., 2015; Cerqua, 2017), where most of the statistical analyses have been carried out, the effects seem much more moderate . As Zielinski and Botero (2019) have noted, some anecdotal evidence seems to have led researchers to think that achieving and maintaining a Blue Flag certification is costly (see Blackman et al., 2014), to the point that it might exceed the economic benefits (see McKenna et al., 2011). However, Pencarelli et al. (2016) report that most of the 2012 Italian recipients of the Blue Flag which had a dedicated budget for it[2] allocated only up to €5K annually.

Using a province-level (*circoscrizioni turistiche*) Italian dataset, Capacci et al. (2015), via a Generalised Method of Moments (GMM) estimation of a demand equation, find that only foreign tourists arrivals are significantly and positively affected by the Blue Flag, while domestic tourists are not affected. However, Cerqua (2017), using different province-level Italian data, finds no evidence of a positive impact of the Blue Flag award on the flow of international tourists and evidence of a positive and significant effect on the flow of domestic tourists, only when the certification comes with a wider sustainability policy. He takes a *reduced-form* approach and estimates the Blue Flag effect via synthetic control methods.

This lack of consensus, even within the same award and country, is in part due to the severe issue of self-selection affecting these environmental awards and eco-labels, for which application is voluntary (see Blackman et al., 2014). Crucially, the lack of information on those who "almost" won and those who "just" won an award – based on some objective criteria – prevents the researcher to compare the most similar units across winners and losers. Moreover, the Blue Flag is awarded through a point-based system, a design well-suited for policy evaluation studies. However, in the Blue Flag case, the NGO operating the programme has never released to researchers[3] neither this information nor information on

[2] 54% of the Italian winners of the Blue Flag in 2012 had a dedicated budget for it.

[3] not even in anonymised version

which municipalities applied and did not win. This implies that the observed non-winners of the Blue Flag include both municipalities that never applied for the certification and those that applied but failed to obtain it. These shortcomings in the data, together with the fact that winners and non-winners are different from each other in important ways, make the construction of a counter-factual scenario, on which the impact evaluation estimation is based, rather difficult and assumption-heavy. Capacci et al. (2015) include in their dataset all the Italian coastal provinces, conditioning their tourism demand variables on a number of covariates that the literature has identified as determinants of tourism demand. They assume that by conditioning on those variables they are able to identify the tourism demand function and hence how tourism demand is affected by the award of a Blue Flag. Blackman et al. (2014) uses propensity score matching (PSM) to build a *control group* (of non-winners) which is then compared with the *treated group*. However, as King and Nielsen (2016) show, PSM often aggravate imbalance, inefficiency, model dependence, and bias, instead of improving on them. Finally, Cerqua (2017), starting from a dataset on tourism flows of all the 164 Italian coastal provinces, selects as the *treated group* 20 provinces that received an extra Blue Flag[4] between 2008-2012, while constructing the *control group* by creating synthetic non-awarded provinces. These are 150 convex combinations of non-awarded provinces with both similar characteristics and pre-award tourism history.

Because of the above-mentioned issues with constructing a valid control group, I construct my estimates via pooled event study analysis (see Cengiz et al., 2019), which relies only on information about the treated units. The idea behind this method is to define a time window around an event of interest (for example, from -2 years from the event to +2 years) and to use the observations outside of this window as a counter-factual for what is happening inside of it. Instead of assuming that the treated group and the control group are identical (in absence of the event/treatment), it assumes that the periods inside the window are identical to the periods outside of it (but for the event/treatment). This assumption is strengthened by the use of time and individual fixed effects, which control respectively for year-specific macroeconomic shocks and individual characteristics.

The other key choice the researcher has to make is what to define as 'treatment'. In the Blue Flag case, the choice is non-trivial, as the award, once won, might or might not be renewed each year, depending on whether the environmental standards are maintained. This could imply that consumers respond differently to a municipality being certified the first time versus one being re-certified for the n-th time, or even one re-obtaining a Blue Flag after losing it. However, this heterogeneity has not been recognised in the literature leading to treatment groups pooling some or all of these cases and hence potentially failing to capture the effect where it has occurred. I argue that the Blue Flag is mostly effective when first awarded, hence focusing on first-time winners. This also allows me to avoid this source of heterogeneity that pooling different types of Blue Flag implies.

My paper's contribution to the literature is threefold. First, by using pooled event study analysis, it avoids the self-selection issue affecting the causal estimation of the effects of interest. Second, it uses data on a more disaggregated level compared to other developed-

[4]without having received an additional one in the previous year

country studies. Third, it looks at balance sheet data, which is useful in in exploring the cost-effectiveness of the Blue Flag certification.

The rest of the paper is structured as follows. Section 2 gives a summary on what the Blue Flag certification is and what are its requirements. Section 3 reports what data is used and the sources from which they were taken. Section 4 explains the identification strategy and the regression specification employed. Section 5 presents the results, while Section 6 summarises the findings and derives some policy implications.

Figure 1: *Blue Flags in Italy, 2018. Map's source: carto.com.*

2 The Blue Flag Programme

The Blue Flag is a programme operated by the Foundation for Environmental Education (FEE), based in Copenhagen. It started as a European programme in 1987 and was then extended to extra-European countries in 2001, so that it is now operating in 49 countries. It is a label awarded to the beaches, marinas and eco-boats that meet a number of environmental, educational, safety- and accessibility-related criteria. In 2017, in Italy, beaches from 230 municipalities received the award. The criteria are verified each year and thus have to be maintained over time to retain the Blue Flag. The mission of Blue Flag is (FEE, 2017):(i) promote and participate in environmental education programmes for the users of beaches, marinas and eco-tourism boats; (ii) implement sound safety and environmental management systems; (iii) monitor environmental conditions to reduce the impact of human activity at the beaches, marinas and eco-tourism boats; (iv) commit to partnerships and collaborative action to promote the sustainable development of tourism. In 2017, 4423 Beaches, Marinas and Eco-tourism Boats in 49 countries (FEE, 2017) featured the Blue Flag, making one of the most successful eco-label worldwide.

3 Data

Several sources are used to construct the dataset. The tourism capacity data between 2002-2016 were taken from ISTAT's website (*Capacity of tourist accommodation establishments*). The municipal balance sheet data for the years 2004-2014[5] were kindly provided by Openbilanci.it, a DEPP and Openpolis project, cofunded by the European Union. Balance sheet information relative to the years after 2014 was not used due to radical changes in municipal accounting system. The Blue Flag data, was partly shared by Capacci et al. (2015), partly scraped from FEE's website[6], the NGO behind this certification. It ranges between 2000 and 2016 and the beach level data have been aggregated at the municipality-level. Table 1 presents the outcome variables used in the analysis.

All the dependent variables are transformed via inverse hyperbolic sine (IHS) transformation. The IHS transformation has the same properties and can be interpreted in the same way as a log transformation. This is because it approximates the logarithmic transformation but for very small values, as shown below:

$$IHS(Y) = log(Y + (Y^2 + 1)^{1/2}) \approx log(2Y) = log(2) + log(Y) \tag{1}$$

The advantage of the log transformation would that it is defined at $Y = 0$, while still allowing for interpreting the regression coefficients in Section 5 as semi-elasticities.

Table 1: *Outcome variables descriptions*

Variable name	**Description**
Tourism services supply	
Total beds	Total beds available in collective accommodations
Hotel beds	Hotel beds available
Other beds	Extra-hotel beds available
Municipal Balance Sheet	
Total municipal revenues	Municipality's revenues from taxes and other sources
Coast-renting revenues	Revenue from renting state-owned properties
Physical capital	Investments in physical capital

[5]Balance sheets are always relative to the past year, so this dataset covers the municipal revenues and expenditures between 2003-2014

[6]www.bandierablu.org

4 Methodology

4.1 Identification

The Blue Flag is a voluntary certification and this makes evaluating its impact challenging. As any voluntary policy or award, it is affected by a problem of *self-section*, in which the winners of the award are not picked at random within the population of interest, but are in fact a sub-sample of the population with specific characteristics. It is those specific characteristics that allow those municipalities (or units, more in general) to receive the award. A naive comparison between the mean outcomes of the two groups of winners (or *treated*) and non-winners (or *non-treated*) would therefore lead to a biased estimate. In principle, this issue could be ameliorated by comparing those municipalities which applied for the Blue Flag and did not get it with those which got it. Indeed, those applying and failing are more likely to be observationally similar to the winners compared to those which did not even apply. This is true under the assumptions that both applicants and non-applicants are aware of the Blue Flag, of its requirements and find it desirable – reasonable assumptions given the long-standing popularity of this beach award. Furthermore, an even stronger identification of the causal impact of the Blue Flag could be provided by having access to the data on all the criteria based on which the municipalities are judged. This would allow the researcher to compare in a regression discontinuity design framework those municipalities that *almost* won the award with those that won it. Given that the FEE does not release any of this information neither to the public nor to researchers, other identification strategies must be adopted.

Figure 2 shows how different the treated and non-treated groups are. It plots a histogram of the probability of receiving the first Blue Flag for the never-certified and ever-certified[7] municipalities between 2002 and 2016, conditional on several census characteristics. Indeed, the two distributions are very dissimilar, the never-certified one being strongly skewed to the left while the ever-certified one being rather uniform. The issue is so severe that for a share of the treated municipalities there is no "comparable" non-treated one with whom comparing them. Because of this fundamental issue of constructing a valid control group, I construct my estimates using a pooled event study analysis (see Cengiz et al., 2019), which relies only on information about the treated units. The idea behind this method is to define a time window around an event of interest – for example, from -2 years from the event to +2 years after the event – and to use the observations outside of this window as a counterfactual for what is happening inside. Instead of assuming that the treated group and the control group are identical (in absence of the event/treatment), it assumes that the periods inside the window are identical to the periods outside of it (but for the event/treatment). This assumption is further strengthened by the use of time and individual fixed effects, which control respectively for year-specific macroeconomic shocks and constant individual heterogeneity. The pooled event study approach is a difference-in-differences approach. It consists in (i) comparing the observations inside the event window with the ones outside via period-specific dummy variables and (ii) taking the difference between the dummies'

[7] i.e. those receiving at least one Blue Flag in the sample

estimated parameters and the parameter associated with a chosen reference period, which must be inside the window and antecedent to the event. Because the pooled event study is a difference-in-differences, the parallel trend assumption must hold in order to identify the causal effect of the Blue Flag: absent of the Blue Flag certification, the outcome variable of interest in treated and untreated municipalities would move in parallel. A further required assumption is no heterogeneity in the effects across municipalities.

This approach avoids issue related to the use of propensity score matching, which have recently emerged in the literature (King and Nielsen, 2016). Blackman et al. (2014) uses propensity score matching (PSM) to build a *control group* of Blue Flag non-winners which is then compared with the *treated group*. However, as King and Nielsen (2016) show, propensity score matching often makes the comparison worse instead of better, by aggravating imbalance, inefficiency, model dependence, and bias. It also avoids the strong assumptions implied by a structural approach as in Capacci et al. (2015). (Capacci et al., 2015) include in their dataset all the Italian coastal provinces, conditioning their tourism demand variables on a number of covariates that the literature has identified as determinants of tourism demand. The causal interpretation of their estimates relies on the assumption that the tourism demand function is correctly specified.

4.2 Empirical Model

In order to study the effect of a Blue Flag certification over a range of municipal-level variables, I take a pooled event-study approach, following (Cengiz et al., 2019). I do this by choosing a a 5-year event window ranging between $[-3, 1]$ in annualised time, where $\tau = 0$ is the year when the Blue Flag is awarded, $\tau = -3$ is three years before the event and $\tau = 1$ is the year after. In the main specification, I look at how receiving a Blue Flag affects tourism services supply and the municipality's balance sheet, in particular during the award year and the following one. When the dependent variable is demand-side, I model the Blue Flag effect as temporary. In particular, I do this by estimating the following regression equation:

$$Y_{i,t} = \sum_{\tau=-3}^{1} \alpha_\tau I_{it}^\tau + \mu_i + \rho_t + u_{it} \tag{2}$$

where $Y_{i,t}$ is an outcome variable, while I_{it}^τ is an indicator variable which equals 1 if the Blue Flag was awarded τ years from calendar year t to municipality i, and 0 otherwise. I also control for both municipality, μ_i, and year, μ_i, fixed effects. After estimating Equation 2, I calculate the (percentage) change of the outcome variable between period -2 and period τ by normalising to α_{-2} the other α_τ coefficients, i.e. by subtracting $\alpha_\tau - \alpha_{-2}$. I choose $\tau = -2$ as the reference year in order to check for anticipation effects in period $\tau = -1$.

4.3 Interpretation

The event of interest is the first Blue Flag certification a municipality wins for one of its beaches. It is expected to increase the economic activity in the associated municipality by

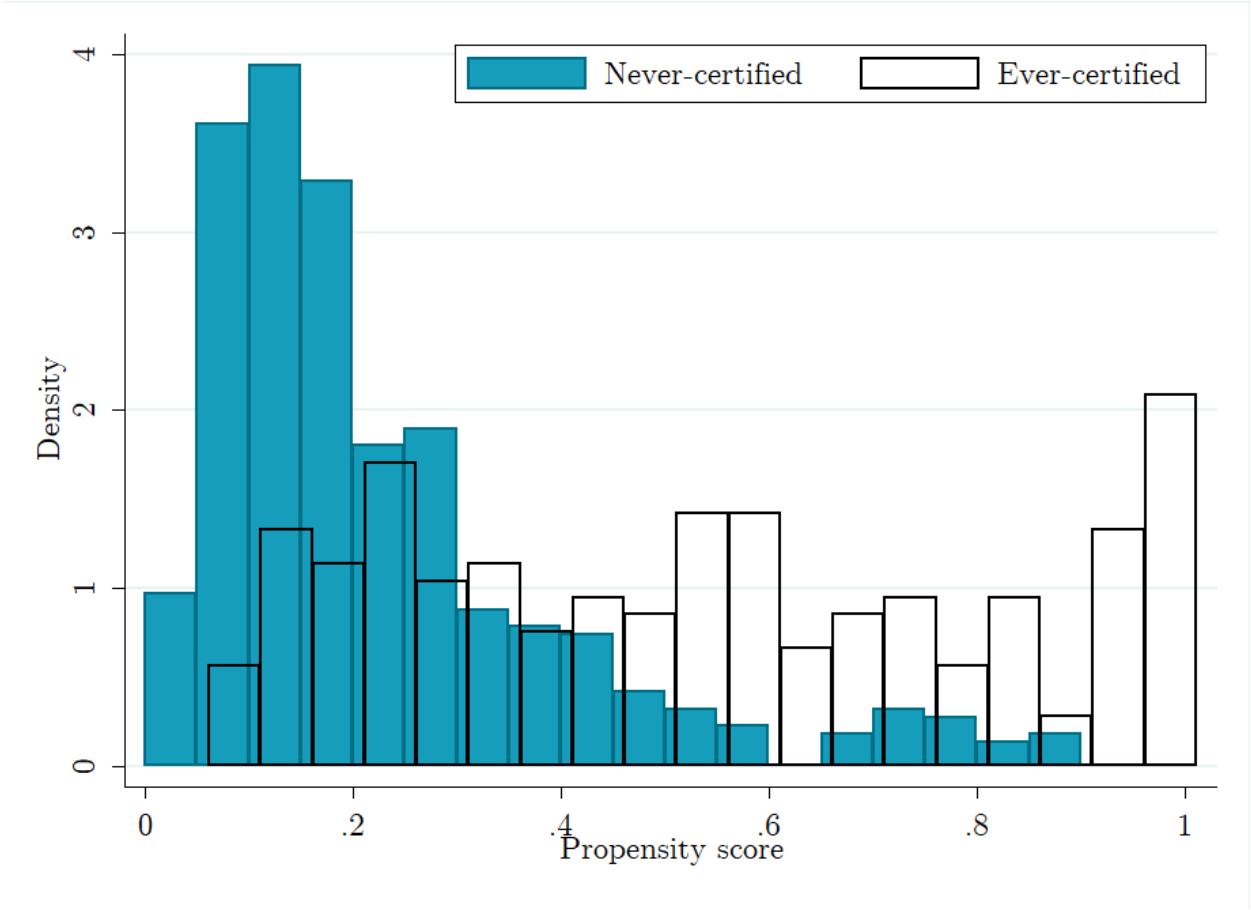

Figure 2: *Distribution of logit propensity scores for never-certified (blue bars) and ever-certified (empty bars) municipalities.*

boosting tourism. A significant increase of the dependent variable at $\tau \geqslant 0$ is evidence that the certification had an effect on it and non-zero effects at $\tau \leqslant 0$ might reveal pre-award patterns, such as anticipation effects and pre-existing positive trends. A casual interpretation of my estimates relies on the assumption that, conditional on municipality characteristics and common macroeconomic shocks, had the Blue Flag not been awarded, the outcome variable would have not significantly changed.

5 Results

5.1 Temporary Effect

Figure 3 shows the temporary impact of being awarded a Blue Flag on municipal revenues. As reported in Table 2, column 1, winning the Blue Flag for the first time increases the revenues of the recipient municipality in the award year by 10% and the coefficient is statistically significant at 1%. A positive and statistically significant coefficient in period $\tau = -3$ could

Table 2: ***Temporary effects of a Blue Flag certification.***

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Municipal Balance Sheet			*Supply of Beds in Collective Accommodations*			
Period	Revenues	Capital	Coastal Rent	Hotel	Extra-Hotel	Total	3+ Stars
-3	.085	.208	.910	-.015	.010	.001	-.033
	(.039)	(.130)	(.715)	(.036)	(.087)	(.044)	(.032)
	[.029]	[.111]	[.203]	[.679]	[.892]	[.988]	[.302]
-2	0	0	0	0	0	0	0
-1	.060	.209	1.863	.035	.047	.053	.109
	(.037)	(.124)	(.680)	(.033)	(.081)	(.041)	(.072)
	[.103]	[.092]	[.006]	[.292]	[.560]	[.197]	[.133]
0	.101	.354	1.596	.021	-.049	.036	.105
	(.037)	(.124)	(.681)	(.033)	(.081)	(.041)	(.095)
	[.006]	[.0046]	[.019]	[.525]	[.548]	[.376]	[.272]
1	.024	-.005	1.149	.022	.031	.041	.073
	(.039)	(.132)	(.725)	(.034)	(.083)	(.042)	(.102)
	[.539]	[.968]	[.113]	[.511]	[.705]	[.320]	[.474]

Note: The table reports the effect of a first Blue Flag certification on municipal balance sheets and tourist accommodation supply. Standard errors in parentheses, p-values in brackets.

mean that the Blue Flag is particularly sought after by mayors as a way to promote the municipality following a decrease in revenues.

As Figure 4 shows, the supply of tourist accommodation does not respond to the Blue Flag certification. Indeed, the coefficients associated with $\tau = 0$ and $\tau = 1$ are not significantly different from zero when the dependent variable is, respectively, *number of hotel beds*, the *number of extra-hotel beds* (including AirBnB's, residences, campings, etc.) the *total number of tourist accommodation beds*, as columns 4 and 5 Table 2 show. The effect remains not statistically different from zero even when looking at the aggregate number of collective tourist accommodation's beds (Table 2, column 6. Given that the award does not significantly impact the overall *quantity* of the supply of accommodations, I study whether it has an affect on the *quality* of such accommodations. I do that by assessing the impact of the Blue Flag on the number of beds in hotels with 3 stars or more and I do not find evidence that accommodation supply is temporarily affected. However, it is not clear whether such supply variables should be considered rigid and hence the award impact on them modelled as permanent. This alternative specification is presented in Section 5.2.

Given the above results, the Blue Flag award seem to temporarily increase demand for

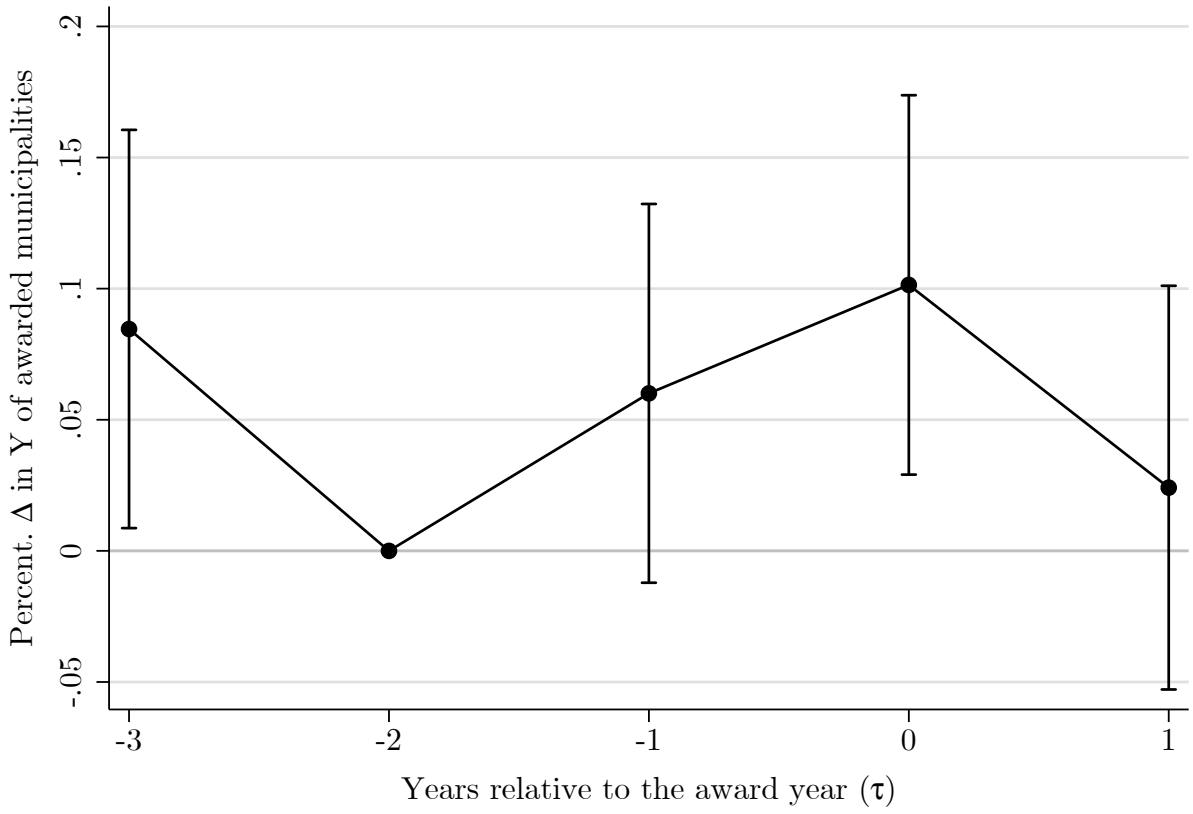

Figure 3: *Temporary impact of being awarded a Blue Flag on municipal revenues.*

tourism, as capture by municipal revenues, although not sufficiently to trigger an increase in supply too, which would be reflected in a positive change in the size of the accommodation industry. This result does not seem to hold in developing countries (see Blackman et al., 2014), where land available for construction is abundant and the tourism sector has a wider margin for growth.

In accordance with the literature (see Zielinski and Botero, 2019), my analysis of municipal balance sheet data shows that on average municipalities do not already meet the minimum requirements when applying for the Blue Flag, but rather invest in order to obtain it. Indeed, physical capital investments start increasing by +20% (significant at 10%) in the year before the award, peaking at +35% (significant at 1%) in the award year (see Figure 5). This is consistent with Zielinski and Botero (2019) which finds that the Blue Flag is (i) a "trigger of political will" that mayors use to effectively allocate resources and coordinate with the local businesses and expertise and (ii) the opportunity for pushing infrastructural improvements.

Revenues coming from renting municipal coastal or maritime areas (*concessioni demaniali*) for commercial use increase significantly and persistently by $10\% - 20\%$ since the

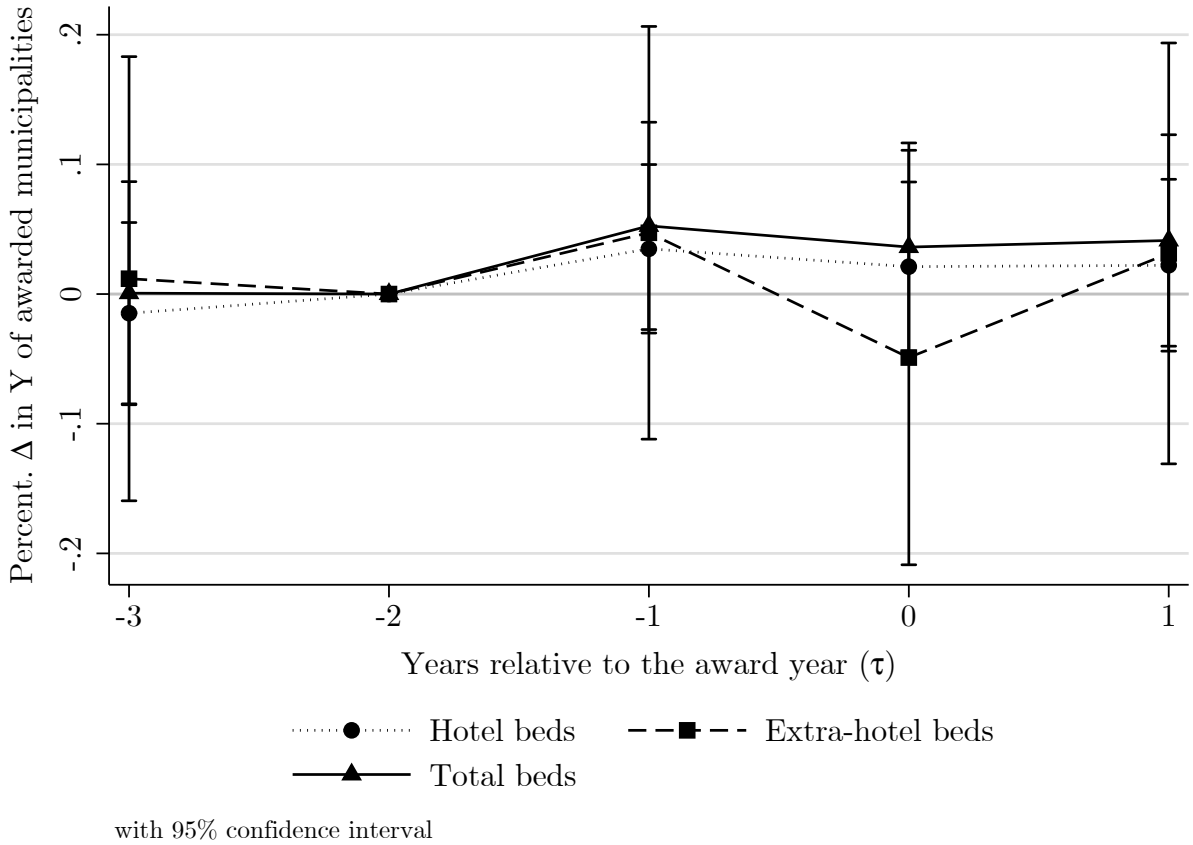

Figure 4: *Temporary impact of being awarded a Blue Flag on supply of collective accommodation.*

year before the award (see Figure 6). Both this effect and the effect on municipal physical capital expenditure start in $\tau = -1$, suggesting the presence of anticipation effects. In other words, this evidence is consistent with winning municipalities investing substantially to meet the Blue Flag's infrastructural requirements while creating positive expectations around the outcome of the Blue Flag application. The private sector responds to this by increasing the commercial exploitation of the coastal areas, which in Italy are property of the government.

In summary, the Blue Flag's effect on municipal revenues is positive and sizeable. In Italy, the municipality budget must be balanced by law (Decreto Legislativo 18th August 2000, n. 267, art. 151), which implies that the Blue Flag's effect on municipal expenses and on municipal revenues are identical (and cancelling each other out). However, if we consider the increase in physical capital expenditure as a long-run investment that will permanently increase the population's welfare, then the net economic effect of a Blue Flag award is positive.

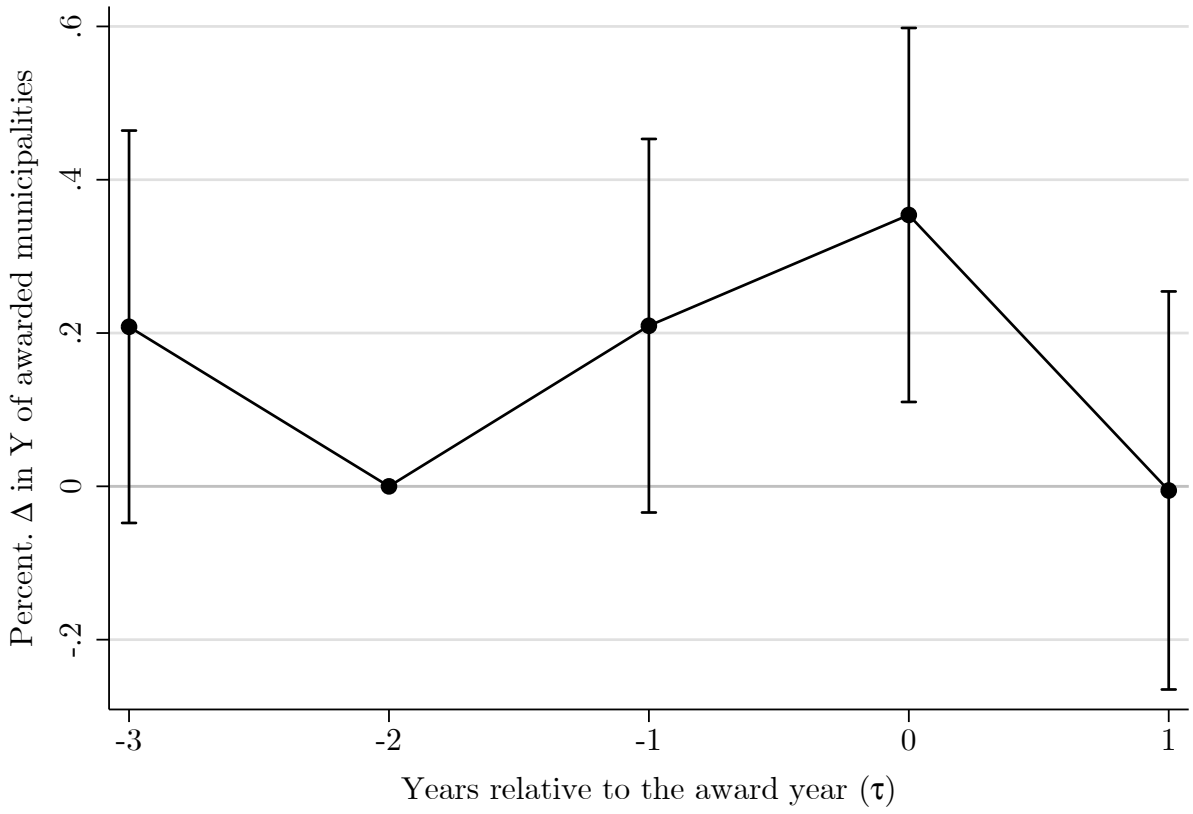

Figure 5: *Temporary impact of being awarded a Blue Flag on physical capital.*

5.2 Permanent Effect

To test the robustness of my estimates, I estimate the same model (see Equation 2), but under the assumption that the effect of a Blue Flag award is permanent, rather than temporary. This is achieved by recoding the $\tau = 1$ dummy variable as equal to 1 when $\tau \geqslant 1$, i.e. in all periods strictly after the event year. It should be noticed that, if an estimated effect is similar when modelled as temporary to when it is modelled as permanent, this is evidence in favour of the temporary-effect specification. Indeed, modelling the effect as temporary means to calculate the counterfactual scenario using both pre- and post-event-window periods, while modelling the effect as permanent means using only the pre-event-window periods. If the effect is modelled as permanent while the true effect is temporary, the estimated effect will similar. However, if the effect is modelled as temporary while the true effect is permanent, the estimates will be biased, given that some periods used to calculate the counterfactual are affected by the award.

I find that modelling the Blue Flag effect as permanent rather than temporary leads to significant coefficients of the same sign and similar value relative to the ones estimated

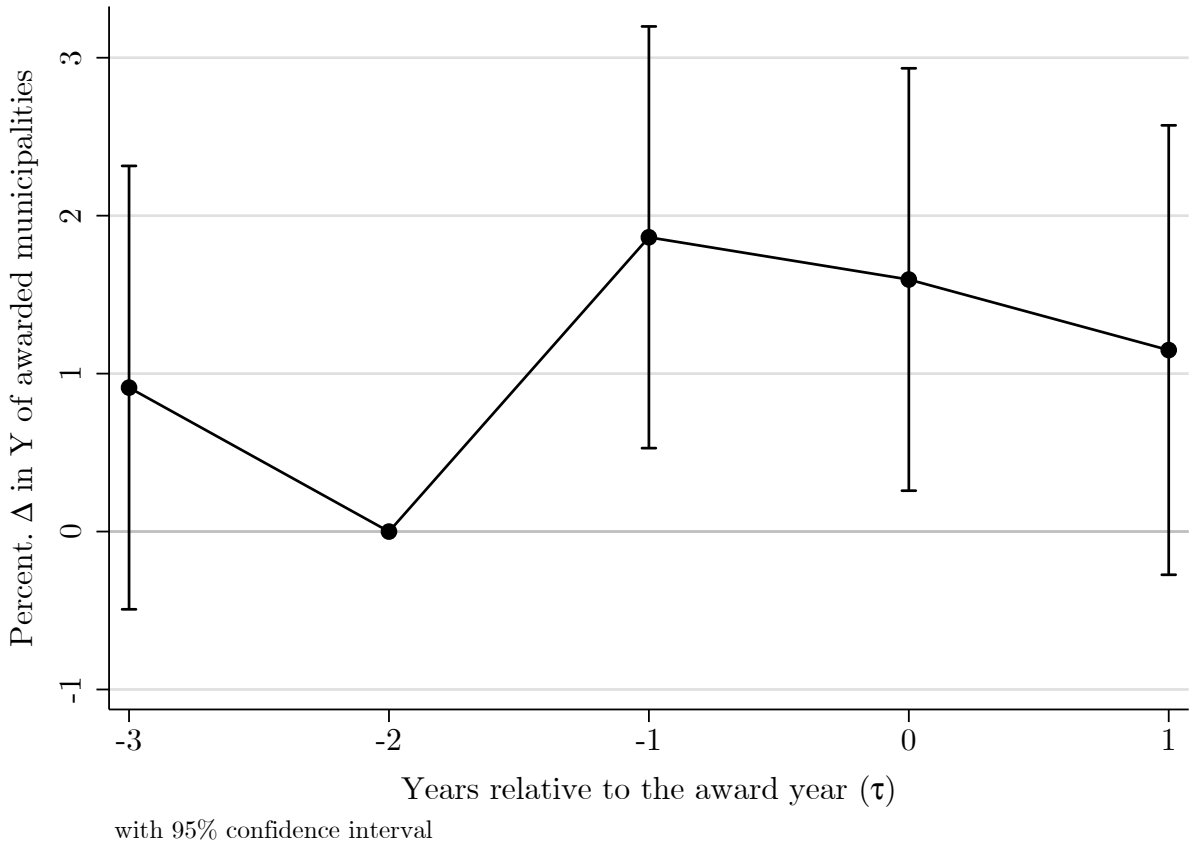

Figure 6: *Temporary impact of being awarded a Blue Flag on municipal revenues from renting coastal public property.*

assuming a temporary effect – with one exception. Figure 7 shows as an example how the main results are robust to this assumption change.

The exception is found when analysing the Blue Flag effect on the supply of beds in hotels with 3 stars or more. As Figure 8 shows, the effect of receiving a Blue Flag for the first time becomes higher and significant at 10% in $\tau = 1$ when it is modelled as permanent. This, as explained above, is (weak) evidence of a permanent effect on this outcome variable, which is consistent with the view that the Blue Flag award increases demand for high quality accommodations and hotels respond by permanently increasing the its supply.

6 Conclusion

Acknowledging the eco-labels' potential in bringing economic and environmental incentives together, a recent literature has developed around them with the aim of testing whether these are effective in practice. The literature evaluating the economic impact of eco-labels has produced mixed results, even when focusing on a single case as the Blue Flag programme.

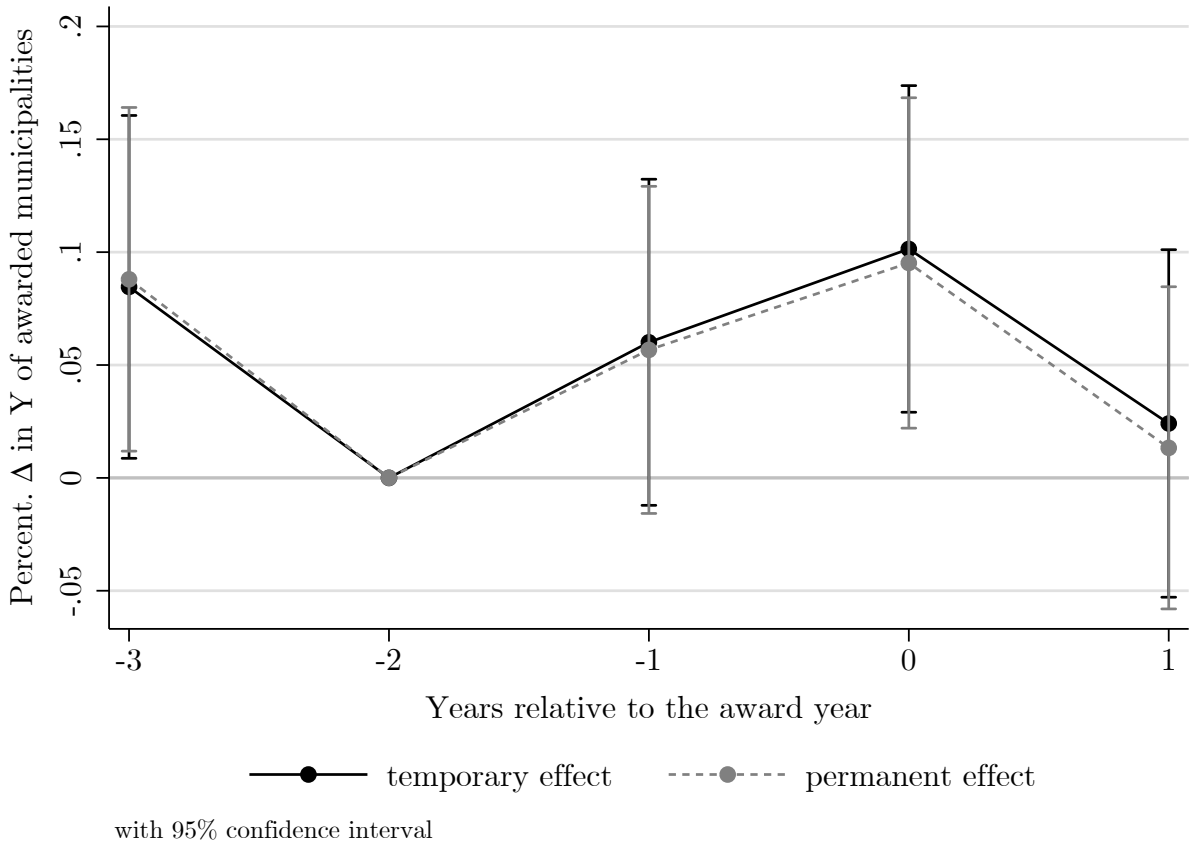

Figure 7: *Temporary and permanent impact of being awarded a Blue Flag on municipal revenues.*

I argue that (i) in absence of detailed data on the applicants, any voluntary certification impact evaluation effort will be sensitive to the econometric assumptions made and (ii) that the Blue Flag literature has wrongly ignored the possibility that the certification might be mainly effecting the first time it is awarded. Therefore, (i) I take an event-study approach, allowing me to avoid the issue of constructing a control group from non-winners and (ii) I focus on the effect of being assigned a Blue Flag for the first time. I find that the effect on municipal revenues of being assigned a Blue Flag for the first time is positive and significant, while I find no evidence that they experience an increase in collective tourist accommodation supply. My findings also provide further evidence (see Creo and Fraboni, 2011; Pencarelli et al., 2016; Cerqua, 2017) not only that the Blue Flag award gives mayors an opportunity to promote and enact environmentally-conscious infrastructural improvements, but also that they are successful in exploiting it, as evidenced by the increased spending on physical capital.

This is the first study to my best knowledge to provide evidence that municipal revenues are positively affected by a Blue Flag certification. Moreover, if we consider the infrastructural investments required by the Blue Flag as a public good, not a dead-weight cost, then the effect of a fist-time Blue Flag certification on municipal profits is positive (see Pencarelli

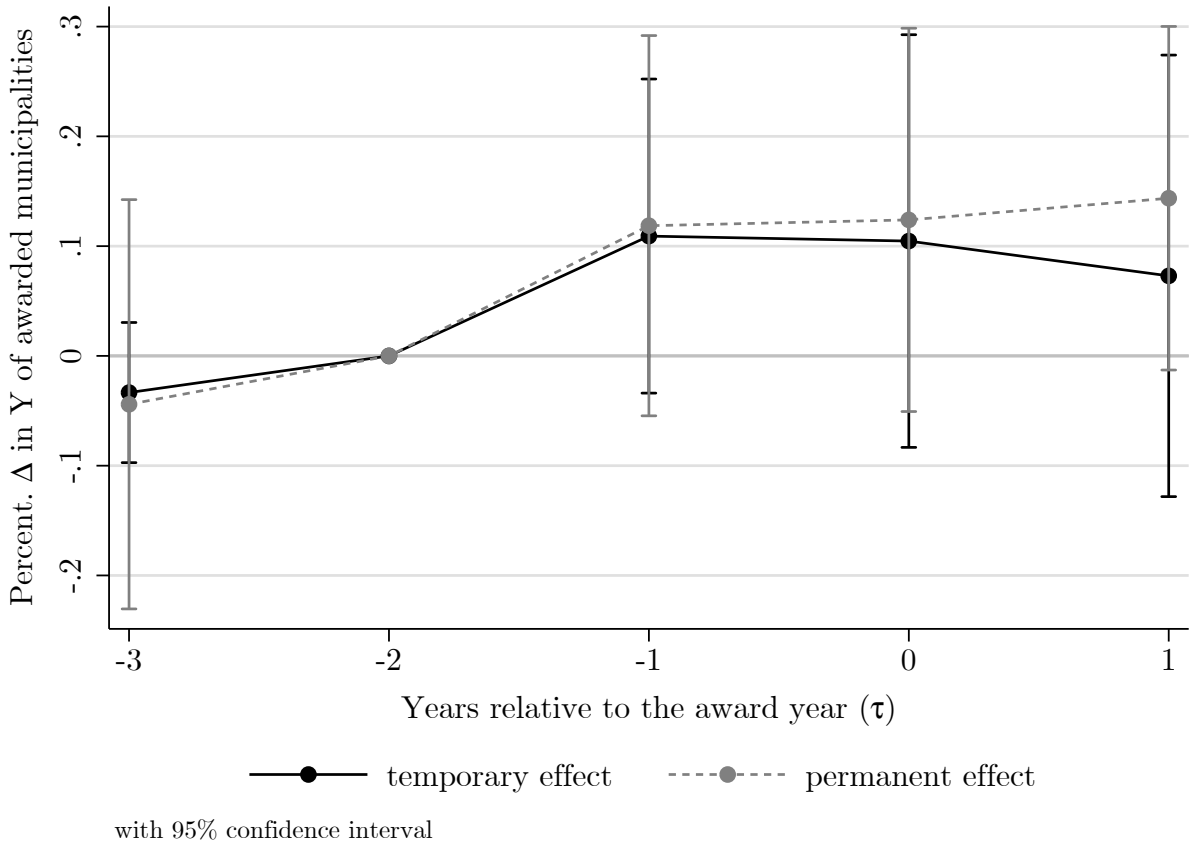

Figure 8: *Temporary and permanent impact of being awarded a Blue Flag on the supply of beds in hotels with 3 stars or more.*

et al., 2016).

References

Blackman, A., Naranjo, M. A., Robalino, J., Alpízar, F., and Rivera, J. (2014). Does tourism eco-certification pay? costa rica's blue flag program. *World Development*, 58:41 – 52. [Cited on pages 31, 32, 36, and 39.]

Blanco, E., Rey-Maquieira, J., and Lozano, J. (2009). Economic incentives for tourism firms to undertake voluntary environmental management. *Tourism Management*, 30(1):112 – 122. [Cited on page 31.]

Buckley, R. (2002). Tourism ecolabels. *Annals of Tourism Research*, 29(1):183 – 208. [Cited on page 30.]

Capacci, S., Scorcu, A. E., and Vici, L. (2015). Seaside tourism and eco-labels: The economic impact of blue flags. *Tourism Management*, 47:88 – 96. [Cited on pages 31, 32, 34, and 36.]

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The Effect of Minimum Wages on Low-Wage Jobs*. *The Quarterly Journal of Economics*, 134(3):1405–1454. [Cited on pages 32, 35, and 36.]

Cerqua, A. (2017). The signalling effect of eco-labels in modern coastal tourism. *Journal of Sustainable Tourism*, 25(8):1159–1180. [Cited on pages 31, 32, and 43.]

Creo, C. and Fraboni, C. (2011). Awards for the Sustainable Management of Coastal Tourism Destinations: The Example of the Blue Flag Program. *Journal of Coastal Research*, 2011(10061):378 – 381. [Cited on pages 31 and 43.]

FEE (2017). Blue flag programme. [www.blueflag.global; Visited on 25-October-2019]. [Cited on page 33.]

Jha, V., Vossenaar, R., and Zarrilli, S. (1997). *Eco-labelling and international trade*. Springer. [Cited on page 30.]

King, G. and Nielsen, R. (2016). Why propensity scores should not be used for matching. *Political Analysis*, pages 1–20. [Cited on pages 32 and 36.]

McKenna, J., Williams, A. T., and Cooper, J. A. G. (2011). Blue flag or red herring: Do beach awards encourage the public to visit beaches? *Tourism Management*, 32(3):576 – 588. [Cited on pages 30 and 31.]

Pencarelli, T., Splendiani, S., and Fraboni, C. (2016). Enhancement of the "blue flag" eco-label in italy: an empirical analysis. *Anatolia*, 27(1):28–37. [Cited on pages 31 and 43.]

Thøgersen, J., Haugaard, P., and Olesen, A. (2010). Consumer responses to ecolabels. *European Journal of Marketing*, 44(11/12):1787–1810. [Cited on page 30.]

Zielinski, S. and Botero, C. (2015). Are eco-labels sustainable? beach certification schemes in latin america and the caribbean. *Journal of Sustainable Tourism*, 23(10):1550–1572. [Cited on page 31.]

Zielinski, S. and Botero, C. M. (2019). Myths, misconceptions and the true value of blue flag. *Ocean & Coastal Management*, 174:15 – 24. [Cited on pages 31 and 39.]