Stuart Taylor Jones

Inclusivity, Enrollment, and Success in Marine Stewardship Council (MSC) Seafood Sustainability Certification Programs

A thesis submitted in partial fulfillment of the requirements

for the degree of:

Master of Marine Affairs

University of Washington

2021

Committee:

Sunny Jardine

Edward Allison

Yoshitaka Ota

Program Authorized to Offer Degree:

School for Marine and Environmental Affairs

University of Washington

Abstract

Inclusivity, Enrollment, and Success in Marine Stewardship Council (MSC) Seafood Sustainability Certification Programs

Stuart Taylor Jones

Chair of the Supervisory Committee:
Sunny L. Jardine
School for Marine and Environmental Affairs

In the past twenty years, seafood sustainability certifications have emerged as a heralded alternative to traditional government-based fisheries management. However, certifying organizations such as the Marine Stewardship Council have often struggled to reach fisheries in the Global South, which are likely among those most in need of improved management. Here, we examine socio-economic factors predicting enrollment and withdrawal from Marine Stewardship Council certifications, and differences in retention and stock assessment outcomes between typical and atypical MSC fisheries. A random forest model identified cost-associated variables, such as existing good governance and high regional biomass, as the most predictive of a fishery's propensity to enroll. However, atypical MSC fisheries, while less likely to enroll, were not broadly less likely to succeed when they did. Withdrawal propensity did not correlate with enrollment propensity, and was predominantly predicted by benefit-associated variables, such as high price, tonnage, and greater international trade. I therefore attribute the documented lack of MSC fisheries in the global South to asymmetric information around the enrollment process. Fisheries likely do not possess good information about what benefits, such as a price premium, to expect from MSC certification, and only those which are already close to certification standards may be willing to take a chance on the program. However, there may be many more fisheries, particularly those which are highly-priced and export-oriented, which could be successful MSC fisheries if they could overcome initial hurdles to reach certification standards. Finally, MSC enrollment was associated with desirable stock assessment outcomes in terms of catch and biomass, but due to data limitations we were only able to evaluate contexts in which it acted as a supplement, rather than an alternative, to government regulations.

Introduction:

As global fishing capacity has grown and total landings have leveled off (FAO, 2018), fisheries managers in the Global North have developed effective methods for controlling effort, preventing stock collapses, and maximizing returns (Hilborn et al., 2020). However, while proven methods such as catch-shares have been implemented disproportionately in developed countries (Jardine and Sanchirico, 2012), most of the world's fishers are in the Global South, where broader institutional and governance factors may inhibit conventional management practices (Allison and Ellis, 2001). As a result, many fisheries in 'developing' countries continue to suffer from greater data limitations, and greater vulnerability of important stocks (Andrew et al., 2007). There is therefore a clear need for alternative solutions and policy tools to manage resources sustainably in contexts which often lack effective regulatory frameworks (Kalfagianni and Pattberg, 2013).

One potential solution is international, incentives-based certifications programs, which can encourage "good behavior", e.g. sustainable capture in wild fisheries, without government regulations. These programs operate principally around labels which convey valuable information about sustainability to seafood consumers in a comprehensible way (Fung et al., 2007). Consumers can then selectively purchase seafood which they believe to be sustainably-harvested (Pascoe et al., 2010; Roheim and Sutinen, 2006). Certifiers are keen to present a certain image to the public (Manach et al., 2020), to create a sustainable reputation which can then be provisioned to fisheries as excludable but non-competitive benefits, or club goods (Prakash and Potoski, 2007). Fisheries are then incentivized to act sustainably so that they can join the 'club' and benefit from this positive reputation through price premiums and consumer preference. To accomplish this, a number of organizations have been created to offer seafood sustainability certifications, including the Marine Stewardship Council (MSC). Formed in 1997 as a partnership between WWF and Unilever, MSC issued its first certification in 2000 and has emerged as the leading seafood sustainability certifier covering over 15% of global catch (Christian et al., 2013; Gulbrandsen, 2009; Manach et al., 2020).

However, MSC's enrollment is dominated by fisheries in already well-regulated countries, and it is unclear whether MSC and similar programs are actually able to reach those fisheries most in need of improvement (Martin et al., 2012; Ponte, 2012). Only 8% of MSC certified fisheries are in developing countries (Stratoudakis et al., 2016), where the need for alternative policy solutions are greatest (Coglianese and Nash, 2001; Kalfagianni and Pattberg, 2013). Existing case studies examining MSC's in the developing world have reported more desirable outcomes (Thomas Travaille et al., 2019) than those located in rich-world countries

(Foley, 2012). Due to the strictly voluntary nature of incentive-based management, evaluating which fisheries participate and why is important to understanding the impacts and limits of this policy tool. A failure to expand into countries with weaker government regulations could limit the overall effectiveness of certifications-based fisheries management.

While MSC has made efforts to be more inclusive of developing-world fisheries, it may face a number of systemic obstacles to expanding enrollment among these kinds of fisheries. First, because the MSC process places the burden to pay on participants themselves, it may be that the program is simply not rewarding enough to entice fisheries which may require substantial changes (Goyert et al., 2010; Martin et al., 2012; Stratoudakis et al., 2016). Price premiums, meant to justify and outweigh the costs of enrollment, have proved elusive, and there is varying evidence as to whether certification has led to significantly better ex-vessel prices (Goyert et al., 2010; Roheim et al., 2011; Sánchez et al., 2020). An analysis of MSCs pre-assessments, which tell prospective fisheries what their chances are and what changes may need to be made, revealed that a majority of fisheries which received 'cautionary issues' ratings (indicating that a few changes may be required) did not proceed with certification (Martin et al., 2012). Meanwhile, because the programs rely on consumer's willingness and ability to pay a premium for more sustainable seafood, there are concerns that fisheries in the Global South are precluded from joining MSC simply because their products are not sold in high-end markets that can pay a sustainability premium (Bellchambers et al., 2016; Oosterveer, 2008; Pérez-Ramírez et al., 2012a). Finally, there are accusations that MSC certifications are fundamentally a matter of protectionism, with developed nations using 'sustainability' as a justification for only buying their own seafood, excluding developing nations which may benefit economically from access to these markets (Kalfagianni and Pattberg, 2013; Oosterveer, 2008).

In this thesis, I examine the impact of a wide variety of socio-economic variables on enrollment in MSC certifications, and continue to follow fisheries throughout the program to assess retention and, where observable, improvements in biomass and other stock assessment variables. In this way, I seek to understand not only the characteristics of a typical MSC fishery, in terms of region, governance, and trade flows, but also the characteristics of a successful MSC fishery. In doing so, I hope to understand who the MSC program serves, and what its benefits are for fisheries management globally.

Background:

Economically, the enrollment decision can be seen primarily as weighing the expected costs of certification against the expected benefits. For costs, under MSC's model, fisheries pay an independent, third-party certifier to assess a number of stock status and management indicators on a 0-100 scale, with 80 and above considered global best practices (Gulbrandsen, 2009) (Gulbrandsen). These evaluations can be very expensive, and must be added to the costs of any harvest restrictions or fishery improvements which must be made in order to achieve suitable indicators (Pérez-Ramírez et al., 2016). Fisheries must achieve a minimum score of 60 (out of a possible 100) for each performance indicator and an average score of 80 or above for each of the three principal categories. For any performance indicator scoring below 80 but above 60, the certifier can assign a condition that, if met, will raise the score to 80 over a specified period of time to a maximum of five years (Christian et al., 2013). Once certified, the client pays the organization an annual fee for the privilege of using their label, which guarantees to consumers that the labelled product was produced in a sustainable way (Gutiérrez et al., 2012). Fees are structured proportionally to the overall tonnage of the fishery (Christian et al., 2013).

In order to justify these costs, fisheries must receive equal or greater benefits in order to remain enrolled, which may come in a number of forms. As described in MSC's original Theory of Change, customers may be willing to pay more for seafood that has been evaluated and certified as sustainable, and this price premium could travel down the supply chain to fishermen and provide an incentive to improve harvesting behavior (Sampson et al., 2015; Uchida et al., 2017). This percent premium multiplied by the initial price and the tonnage makes up the main economic benefit for enrolling fisheries. However, many fisheries, particularly in the developing world, may sell some or all of their product to markets where consumers are unwilling or unable to pay this premium, and this ineligible tonnage must be subtracted from this term. In recent years, MSC's pitch has pivoted to focus on market access (Uchida et al., 2017), and some fisheries may be pursuing greater market share rather than necessarily higher prices. MSC certifications have also allowed cooperatives (Pérez-Ramírez et al., 2012b) and processors (Foley, 2012) to assert greater control over resources. Finally, fisheries may seek MSC's reputation for sustainability in order to shield themselves from further regulation or from public criticism over their business practices (Izquierdo-Peña et al., 2020).

Once the enrollment decision has been undertaken, actual benefits that fisheries experience must continue to outweigh costs for fisheries to remain enrolled. A high rate of voluntary withdrawals may indicate that benefits do not exceed costs under our economic model for a large proportion of fisheries. Conversely, fisheries which do persist with the program are likely those for which the enrollment decision was most economically sound, and which do realize net benefits from their participation. We therefore contend that understanding not only enrollment but retention and eventual improvements are key to understanding which fisheries are benefitting from these certification programs.

Similarly, fisheries whose realized benefits more greatly exceed administrative costs may be more willing and able to take on harvest restrictions and other improvement costs which are the real raison d'etre of the program. Though some certifiers, namely FIPs, evaluate fisheries in a much more holistic way, including in terms of their social responsibility and fair business practices, MSC's goals are defined much more narrowly (Gulbrandsen, 2009). The program's three overall criteria are: 1. Stock Status 2. Environmental Impact 3. Effective Management, all of which should be reflected in the ecological outcomes of target fisheries. The success of MSC programs, then, according to their own criteria, should be measurable in population statistics reported in stock assessments. I was therefore interested in evaluating stock assessment data before and after MSC enrollment in order to determine whether there were in fact observable improvements, and whether these were affected by the socio-economic contexts of enrolled fisheries

Methods:

Data collection:

In order to explore the inclusivity and success of MSC programs at an international level, this thesis needed large quantities of socio-economic data covering broad geographic and temporal ranges and including both certified and uncertified fisheries. These data were drawn mainly from 7 publicly-available global datasets: FishSource, World Governance Index, Human Development Index, RAM Legacy Stock Assessment database, FishStatJ, FishBase, and MSC's own website. Data were collected from online repositories in June 2020, both by hand and by using data-scraping tools such as ParseHub, and have not been updated to reflect recent updates between the collection date and the submission date of this document. These data were collected at five levels; stock, country, region, product, and species. All data were collected as a panel for each year between 1997, three years before the first MSC certification, and 2020, or the last year for which data were available at that time (Table 1). The only exception were species' life-history traits, which were not considered to vary significantly over decadal time scales.

Variable	Enrollment Mo	Withdrawal	Outcomes M	Source	Construction

K	x	x		Fishbase	Average of all observations listed on fishba Bertalanffy growth curve coefficient, eg. rat species. For species without listed values, average of other species in that genus, and none then the average of other species in t
AfM	x	x		Fishbase	Average of all observations listed on fishba maturity of a species, in years. For species values, I substituted the average of other s genus, and if there were none then the ave species in that family
Fec	x	x		Fishbase	Average of all observations listed on fishba absolute fecundity of a species, or the total it can produce over its lifetime. For species values, I substituted the average of other s genus, and if there were none then the ave species in that family
M	x	x		Fishbase	Average of all observations listed on fishba mortality, a model value which shows the to expected to die of natural causes in a year expressed as a proportion of standing popu year. Values may exceed 1 for species whi several times in the course of a year or wh short adult life stage. For species without li substituted the average of other species in if there were none then the average of othe family
realTrade	x	x		FishstatJ	Most recent reported value (2017) for total products derived from a given species, or g similar species (i.e. salmons, cold-water sh mackerels), traded internationally in a year
realValue	x	x		FishstatJ	Most recent reported value (2017) for total all products derived from a given species, o similar species (i.e. salmons, cold-water sh mackerels), traded internationally in a year
realPrice	x	x		FishstatJ	Most recent reported value (2017) for total products divided by total volume of product given species or group of very similar spec internationally in a year
Imports_2017_MT	x	x		FishstatJ	Most recent reported value (2017) for total and exports, and arithmetic and geometric FishstatJ for each country
Exports_2017_MT	x	x		FishstatJ	Most recent reported value (2017) for total and exports, and arithmetic and geometric FishstatJ for each country
Arith_Balance_2017_M	x	x		FishstatJ	Most recent reported value (2017) for total and exports, and arithmetic and geometric FishstatJ for each country

Geom_Balance_2017_l	x	x		FishstatJ	Most recent reported value (2017) for total and exports, and arithmetic and geometric FishstatJ for each country
Imports_2017_kUSD	x	x		FishstatJ	Most recent reported value (2017) for total and exports, and arithmetic and geometric FishstatJ for each country
Exports_2017_kUSD	x	x		FishstatJ	Most recent reported value (2017) for total and exports, and arithmetic and geometric FishstatJ for each country
Arith_Balance_2017_kU	x	x		FishstatJ	Most recent reported value (2017) for total and exports, and arithmetic and geometric FishstatJ for each country
Geom_Balance_2017_l	x	x		FishstatJ	Most recent reported value (2017) for total and exports, and arithmetic and geometric FishstatJ for each country
Biomass			x	RAM Legacy	Reported time series of population biomass year beginning in 1997
Catch			x	RAM Legacy	Reported time series of total catch values f beginning in 1997
Region_Bmsy	x	x		RAM Legacy	Mean of all available RAM values for stock biomass/biomass at MSY. Most recent valu each fishery, but only if that value was from years (2010-2017)
Region_Umsy	x	x		RAM Legacy	Mean of all available RAM values within the for total harvest/harvest at MSY. Most rece used for each fishery, but only if that value 10 years (2010-2017)
R_assess_rate	x	x		RAM Legacy	Percentage of fisheries within each RAM-d which have stock assessments listed in the
C_Assessed_rate	x	x		RAM Legacy	Percentage of fisheries in each country wh assessments listed in the RAM database
Assess.fac	x	x		RAM Legacy	Binary value (1 or 2) indicating if a stock ha assessment listed in RAM
HDI	x	x		World Bank	2017 Human Development Index for each
Accountability	x	x		World Bank	World Bank indexes showing 2018 expert a given government's Accountability, Stability Regulatory Quality, Corruption, and Rule of
Corruption	x	x		World Bank	World Bank indexes showing 2018 expert a given government's Accountability, Stability Regulatory Quality, Corruption, and Rule of
Regulatory Quality	x	x		World Bank	World Bank indexes showing 2018 expert a given government's Accountability, Stability Regulatory Quality, Corruption, and Rule of

Stability	x	x		World Bank	World Bank indexes showing 2018 expert a given government's Accountability, Stability Regulatory Quality, Corruption, and Rule o
Effectiveness	x	x		World Bank	World Bank indexes showing 2018 expert a given government's Accountability, Stability Regulatory Quality, Corruption, and Rule o
Rule of Law	x	x		World Bank	World Bank indexes showing 2018 expert a given government's Accountability, Stability Regulatory Quality, Corruption, and Rule o
MSCcombined		x		MSC.org	Yes/no variable showing whether a certifica combined with another pre-existing certifica
MSCtonnes		x		MSC.org	Tonnage of fish enrolled in a given certifica
Myears.enrolled			x	MSC.org	MSC's start date subtracted from present c

Table 1: Details of variables used in each model

First, to define the world's fisheries, I created a reference list of internationally-traded fisheries from FishSource ("FishSource," n.d.), an online database primarily intended for seafood buyers in developed countries. FishSource lists data by stock, and defines a fishery as each unique combination of stock x country x gear. The database therefore records the number of country x gear combinations targeting each stock, and states which and how many of these fisheries enroll in MSC and FIP programs.

To add specific information on MSC status, MSC certifications listed on FishSource were matched by hand with meta-data from MSC's own websites, ("Fisheries - MSC Fisheries," n.d.). This included the year of enrollment in the program, as well as the certification's current status, such as whether it had been withdrawn by the fishery, suspended by the MSC, or was still in assessment, though these events did not always have listed dates. MSC's website also provided important context data, such as the total landed tonnage enrolled in each certification, and whether it had been combined with another certification since its establishment.

To create a framework to be populated by the remaining context data, I identified for each fishery from FishSource the country, stock, species, product type, and geographic region involved. Additional data were then added to this framework by matching the relevant field between the database and the framework. For example, since each 'fishery', as defined by FishSource, is operated by a single country, we matched each fishery with the appropriate country-level data from the World Bank on the World Governance Index and Human Development Index. The WGI reports six primary numeric indices (Accountability, Corruption, Effectiveness, Regulatory Quality, Rule of Law, Stability) yearly for each UN member state, beginning in 1997. For HDI, only the namesake Human Development Index was used as an indication of country wealth.

To match MSCs with specific stock assessment outcomes, each fishery was also matched with publicly-available fisheries and population data from the RAM legacy database, where available. The RAM database collates data on a wide variety of ecological and biological metrics as available, though because each fishery decides on its own which variables to assess, many more specific variables are inconsistently available. The most commonly collected, and thus the most commonly reported, metrics were simple catch and biomass, followed by the ratio of biomass vs. biomass at maximum sustainable yield (B/Bmsy) and the same ratio for harvest (U/Umsy). I aggregated these two ratio variables across each listed RAM region (ex: US Alaska, US East Coast, South America, etc.) to create a regional context variable which indicated how sustainable fisheries in that region generally are in relation to MSY.

Fishbase was also used to gather life history data on each species targeted, in order to distinguish between target species with 'fast' life histories, such as anchovies, and those with 'slow' life histories, such as orange roughy. For each species, I took the average value reported across published studies for four variables: Natural Mortality (M), von-Bertalanffy growth coefficient (k), Absolute Fecundity (Fec), and Age at first Maturity (AfM). While life history data were not universally available, I was able to use the species' relative taxonomies in order to roughly fill in gaps. For species for which a certain value was not reported, I used the average of all other species in its genus, and for those which had no reported data in their genus, the average value of species in its family.

Trade data on seafood products were also taken from FishStatJ, which is maintained by the FAO, in order to understand how likely certain fishes and certain countries were to be export-oriented. Trade flows in seafood products were totalled by country to estimate, for instance, which countries exported large amounts of seafood, or exported much more seafood than they imported. Trade data were also compiled by species/product (for instance, salmons, anchovies, etc.) in order to determine which products were highly traded or commanded a high price. All data involving a monetary value were collected in USD and controlled for inflation using the consumer price index (CPI).

Enrollment model:

To identify the socio-economic characteristics of a typical MSC fishery, we created a predictive statistical model of enrollment using a technique known as a random forest. The technique is essentially a multi-iteration classification and/or regression tree (CART), which utilize a large number of dependent variables, such as our publicly-available context dataset, to predict a single outcome variable by creating a hierarchical 'flowchart' of binary recursive splits (Breiman et al., 1984). For each 'split', or node, the algorithm sorts the data into two groups based on the value of one of the explanatory variables, chosen in order to maximize the homogeneity of the response variable in each of the two groups. They are robust to multi-collinearity, and substantially relax typical assumptions of regression methods (normality, homoscedascity, etc.), making them broadly useful for a wide variety of datasets and applications (Breiman et al., 1984). This modelling method was preferred to a more conventional logit model, as random forests have been shown in benchmark studies to generate more accurate predictions (Couronné et al., 2018).

Individual decision trees, however, can be unstable and overfit a model to the exact data they are trained on (Hastie et al., 2009). To overcome this, a random forest creates a very large number of decision trees using different subsets of observations and explanatory variables, in order to create more stable and generalizable predictions (Hastie et al., 2009). The random forest begins by randomly selecting two-thirds of the available training data, in a process known as bootstrapping, and then creating a decision tree based on this subset. This process is then repeated many times, creating a 'forest' of decision trees each based on a different random sample of the training data (Varian, 2014). The respective outcomes predicted by each of these individual trees are then aggregated together for each observation to create a better estimator.

In our study, a random forest was 'trained' to predict MSC enrollment on fisheries which did not have any listed assessment in the RAM database, since these could not be evaluated for 'success' in terms of catch and biomass outcomes. The training data included all of the context variables (Table 1), such as country-level data, life history traits, and regional averages of biomass vs MSY, but not stock-level assessment data. These data were taken from the year of enrollment for MSCs, to ensure exogeneity, and from 2017 for non-enrolled fisheries, which was the most recent year for which all data were reported. The algorithm was trained to use this data to discriminate between fisheries which have ever enrolled in MSC from those which have not. Fisheries which have enrolled in FIPs were also excluded from the dataset. To deal with missing data, the random forest used a 'rough fix', which essentially replaces missing data with the median value for that variable. Because the CART algorithm is based on binary splits, this means that when a given explanatory variable is used for a split, observations which are missing that explanatory variable will always remain with the larger group.

Once the random forest was created, the model was then applied to the 'test' data. The algorithm created a 'propensity score' for each assessed fishery, a value between 0 and 1 which represents the percentage of decision trees within the forest in which the fishery was predicted to enroll in MSC. In the context of this study, this propensity score can be considered to represent how 'MSC-like' a fishery looks, in terms of the socio-economic context variables. In order to account for missing values in the test data, we imputed missing context data using the missRanger package, and these imputed values were included when calculating the propensity score. The missRanger package generates actual best-guess values for missing variables based on those variables which are available for that case, including data from the same case in years before and after those which are missing. This is not the same as the 'rough fix' which is endemic to the random forest package and simply defaults cases into the larger group when they are missing the variable used in a particular split. While the percentage of missing data as a proportion of total data is very low, due to the high number of explanatory variables and datasets involved, almost all cases were missing data for at least one variable for some years. Because the algorithm which generates the propensity scores from the test data cannot handle any missing predictors, restricting the analysis to only cases which had no missing variables would have restricted sample size so greatly as to be meaningless.

Withdrawal Model:

To examine the retention aspect of 'success' in MSC programs, I created a second random forest model to predict withdrawals, defined here as fisheries which voluntarily exit the MSC program, thus excluding fisheries whose certifications have been suspended. This model used the same context variables as the enrollment forest, with the addition of MSC-specific variables such as total landed tonnage enrolled in the program and whether the certification was combined with another pre-existing certification. The withdrawal parameters were optimized using the same method detailed above and were ntree=150 , mtry=5, and nodesize=1. Withdrawal propensity was also tested for correlation with enrollment propensity using simple linear regression.

Outcomes Model:

To examine the catch and biomass improvement aspect of success, I used the Generalized Synthetic Control Method (GSCM) to create a synthetic control, or counterfactual, to compare MSC fisheries against (Abadie et al., 2010; Xu, 2017). Synthetic controls are a form of difference-in-difference inference, observational methods in which trends over time in a variable of interest are compared before and after a given treatment to controls which followed similar pre-treatment trends or were otherwise similar to the treated group, in order to determine the effect of the treatment. However, since there is rarely a perfect, identical case to use as a control, synthetic control methods have emerged as an important tool to create counterfactuals which mirror pre-treatment trends of the treated group. Rather than use any existing individual case as a control, a synthetic control combines multiple untreated cases in order to minimize the mean standard pre-treatment error (MSPE), or the difference between the treatment and control data before the actual intervention. By minimizing the MSPE, this synthetic combination should also match what the post-treatment data *would* have looked like in the treatment group. Any difference between the synthetic control and the actual treatment data in the post-treatment period can then be interpreted as the effect of the treatment.

Given MSCs focus on ecological outcomes (Gulbrandsen, 2009), we evaluated biomass and catch data from RAM stock assessments before and after MSC enrollment to examine the outcomes of the program (Figure 1). For each MSC fishery, the GCSM weights the stock assessment data from several unenrolled control stocks to create a synthetic control that matches that MSC fishery's stock assessment data as closely as possible up to the start of the enrollment process. This weighted linear combination of stock assessments is then compared against assessment data from the MSC fishery after treatment, to determine the effect of enrollment on stock status.

Prior to using the GCSM, biomass and catch were normalized by dividing each year by the stock's 1996 value, the year prior to the start of our panel dataset. This ensured that fisheries of different sizes would not be weighted differently when computing summary statistics for the overall GCSM. One attractive feature of the GCSM, relative to the standard synthetic control method, is that it is a non-parametric method, which helps distribute weights more evenly among control fisheries (Cole et al., 2020) and avoid issues with 'corner solutions' (Kuosmanen et al., 2021). I experimented with defining 'treatment' as beginning at the official start of MSC certification as well as 1, 2, 3, and 4 years beforehand, since fisheries can be expected to make changes and undergo improvements at any time in advance of certification in anticipation of a future assessment (Abadie et al., 2010). Treatment year was defined as beginning at three years prior to certification, which was also typically the start of the initial assessment in pre-2012 MSCs[1], which make up almost the entirety of the selected treatment group.

To allow for the best possible comparisons, I selected subsets of both treated MSC and control fisheries (Figure 1). Using the propensity score, I selected a group of candidate control units which have a similar likelihood of enrollment as the MSC treated fisheries, such that observed differences in socio-economic contexts between typical MSC and non-enrolling fisheries could be controlled for. Rather than using the propensity scores to match fisheries directly (King and Nielsen, 2019), we simply narrowed the group of candidate controls to those within the same region of support as the treatment MSC fisheries. For MSCs we selected only fisheries which did not target the same stock as another MSC, which reported the outcome

---

[1] The assessment process seems to have been streamlined in 2012, as fisheries progressed from initial assessment to certification more quickly after this date

statistic of interest, and which had sufficient available data, at least 8 years of data before treatment and at least 8 years after treatment (3 'anticipation years' and 5 enrolled years). Control assessments were selected which had propensity scores higher than the lowest-scored MSC in the treatment group (0.36), which did not target the same stock as any MSC or FIP fisheries, which reported the statistic of interest, and had at least 16 years of data.

Results:

Enrollment:

The predictive model performs well in identifying MSC vs unenrolled fisheries in the test dataset. Propensities scores had a correlation of 0.71 with MSC status (0 or 1), and when binary classification was enforced had a correct classification rate of 82%, including 75% of MSCs and 92% of unenrolled fisheries. Earlier MSC fisheries had generally higher propensity scores while fisheries which have enrolled since 2016 were generally more representative of the rest of the control data (Figure 1).

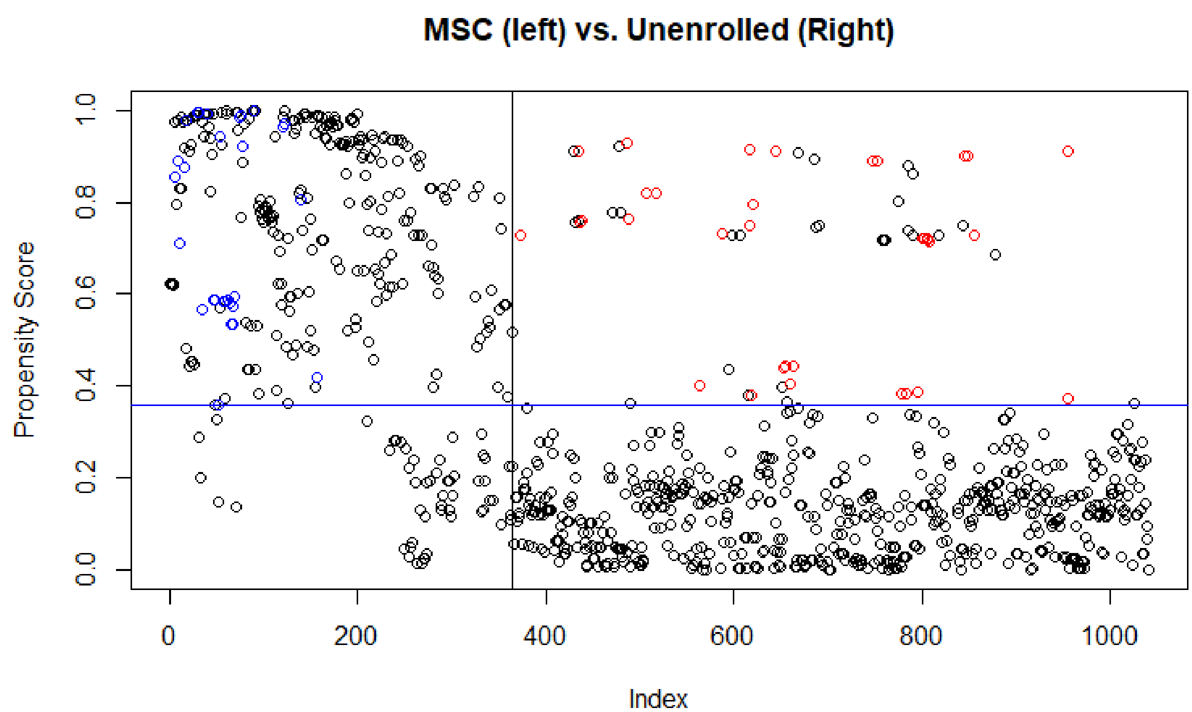

Figure 1: Enrollment Propensity: All assessed fisheries are scored 0 to 1 based on their likelihood of enrolling in MSC. Actual MSC fisheries are on the left, ordered left-to-right chronologically in terms of start of certification, while unenrolled fisheries are on the right. For the GSCM, suitable treatment fisheries are in blue, while suitable control fisheries are in red

To generate its propensity predictions, the random forest model relied heavily on regional biomass and harvest rates, more so than country-level governance (Table 2). Well-regulated countries, such as the USA, targeted a wide variety of regions, both within and beyond their exclusive economic zones; however, MSC fisheries were concentrated in the most pristine regions (US Alaska) while there were few in more historically-exploited areas targeted by the same countries (US East Coast) (Figure 2). Countries with lower governance scores were generally not able to target these pristine areas, and most of their MSCs targeted open-ocean commons (Pacific, Atlantic) (Figure 2). Country-level seafood trade statistics were also important in predicting MSC enrollment, as fisheries countries with high background seafood exports and export-heavy trade balances were more likely to enroll (Table 2). The model considered these variables to be more or less as helpful as governance data (Table 2).

Enrollment				**Withdrawal**			
Variable	Node Purity	Variable	%MSE	Variable	Node Pu	Variable	%MSE
Regional Biomass		Regional Biom	27.8	Regional Harvest		Regional Harvest	
Seafood Exports (Val		Regulatory Qu	14.4	Enrolled Tonnage		Enrolled Tonnage	
Regional Harvest		Rule of Law	12.7	3-year Price Trend		National Trade Bala	
Regulatory Quality		Stability	12.5	Intl Market Price		3-year Price Trend	
Trade Balance (Value		Seafood Expo	11.7	Intl Market Volume		Natural Mortality	
Rule of Law		Government E	11	National Trade Balanc		Intl Market Price	

Table 2: Two metrics of variable importance for data used in the enrollment and withdrawal models. Cost-associated variables are in red, while benefit-associated variables are in green

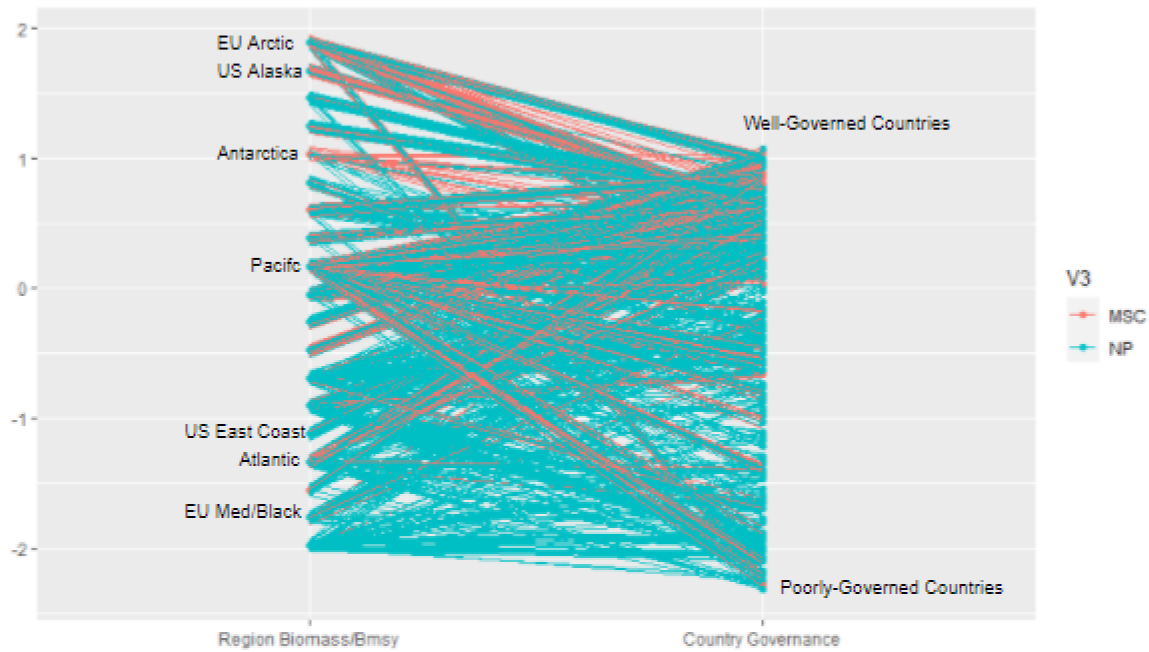

Figure 2: On the left, RAM regions listed top-to-bottom by biomass/MSY, and on the right, countries listed by 'Regulatory Quality' index from the WGI. Each line represents a fishery, fished by the country on the right, targeting the region on the left. MSC fisheries are in red, all other fisheries are in blue.

Withdrawals:

As of April 2021, MSC has issued over 370 certifications over the course of its 21-year history, but over a third of these (130) are no longer active. Of these 130, 114 were voluntarily withdrawn by the fishery itself, with only 16 being suspended by the MSC (msc.org). The withdrawal model was able to distinguish fisheries which voluntarily withdrew from MSC from those which remained with a correct classification rate of 88% (88% for withdrawals and 89% for retained fisheries). While the withdrawal model had access to all of the same data as the enrollment model, it made use of very different variables to generate its predictions (Table 2). Indeed, withdrawal propensity scores did not correlate with enrollment propensity scores (Rsq=0.0014, p=0.21) (Figure 3). In our dataset, withdrawals were found primarily in smaller MSCs tonnage-wise, which targeted a low-priced species, and which were not combined with other MSC certifications.

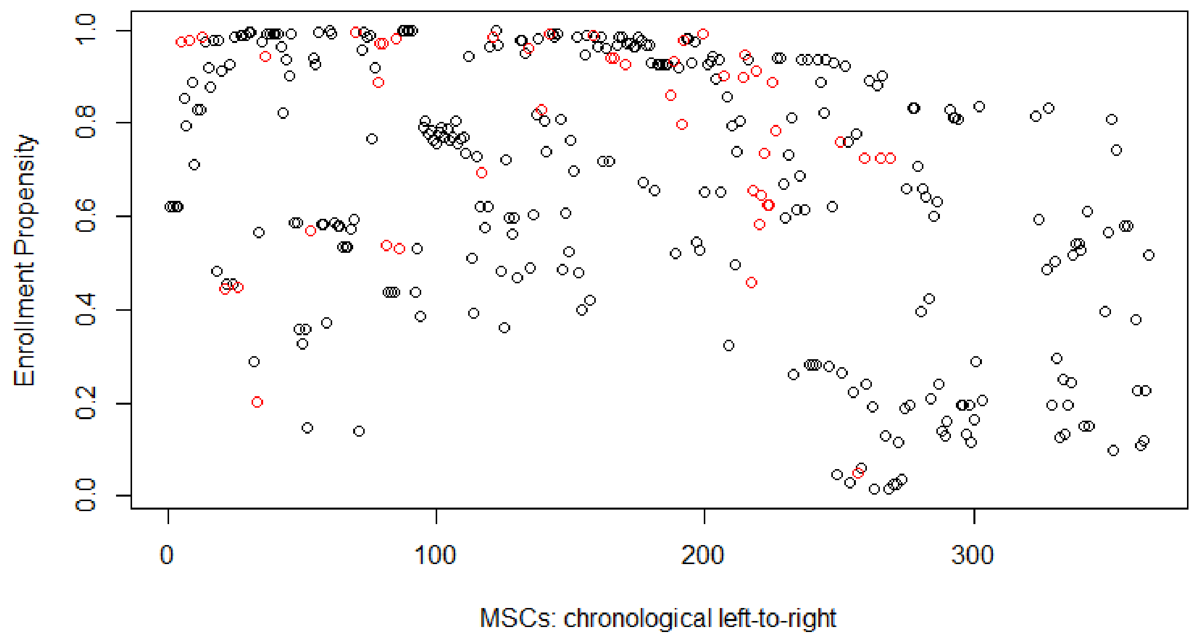

Figure 3: Enrollment propensity scores of MSC fisheries, ordered chronologically left-to-right. Fisheries which voluntarily withdrew are highlighted in red.

Outcomes:

We identified 27 MSC fisheries which were suitable treatment units according to our methods, and these fisheries showed both short- and medium-term improvements in biomass which were attributable to certification. Overall, biomass improvements were greatest in the certification year and the year after, with increases of 30±18% of 1996 biomass in year 3 after treatment and 44±31% of 1996 levels in year four after treatment (Figure 4). While improvements declined after these years, they remained positive and were again significant seven years after treatment, four years after certification (Figure 4). Biomass improvements were consistent whether the MSC treatment fisheries were compared to all controls, or only to high-propensity, MSC-like controls. Catch, meanwhile, remained relatively constant in the treatment fisheries, while the trends in unenrolled fisheries departed based on whether those fisheries were high-propensity. Catch data for synthetic controls constructed from all available unenrolled fisheries moved sharply up after treatment began, but remained relatively constant in high-propensity fisheries (Figure 4). As a result, treated MSC fisheries only showed improvements in catch data relative to all controls.

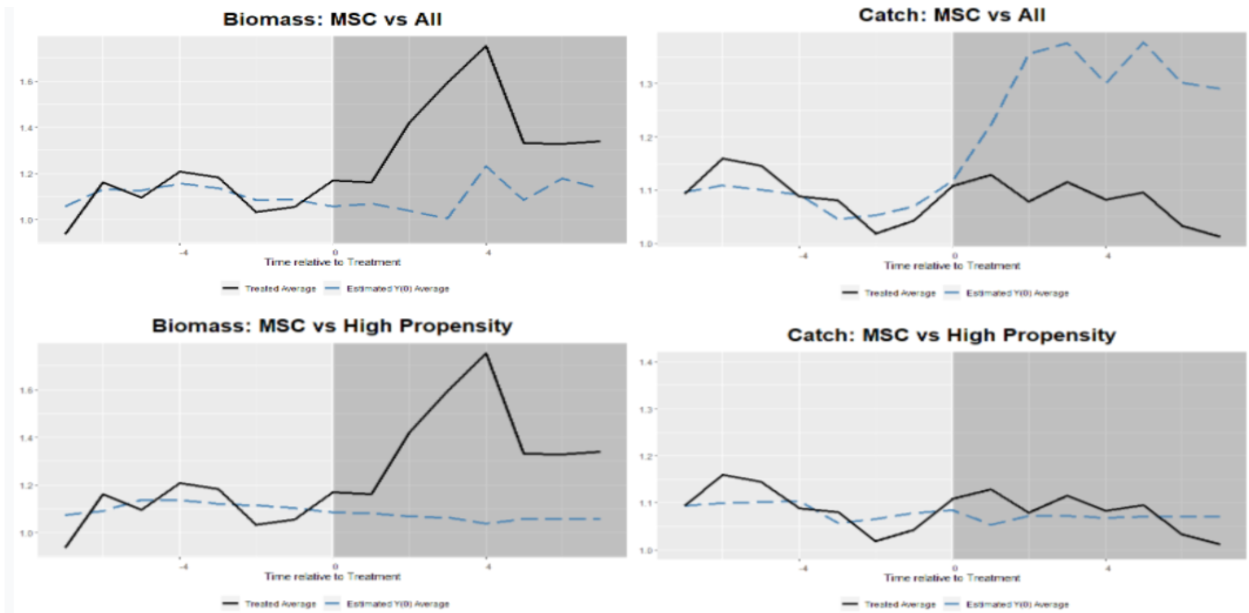

Figure 4: Treated MSC biomass (left) and catch (right) in black vs synthetic controls (dotted blue) constructed using all unenrolled fisheries (top) or only high-propensity, MSC-like controls (bottom). Graphs divided into pre-treatment (light grey) and post-treatment (dark-grey) periods. Treatment is defined as beginning three years before the official certification date

Discussion:

The success of the predictive model in generating accurate propensity scores in the test data demonstrates that MSC fisheries tend to share characteristic socio-economic profiles, specifically related to regional biomass and governance indices. MSC fisheries tend to be located in very pristine fishing regions, and are often fished by countries considered to be well-governed; however, this likely reflects that richer countries have greater access to the distant, polar regions which have the most pristine biomass (Figure 2), as regional biomass is much more predictive of enrollment than governance scores (Table 2). 'While these areas are targeted primarily by well-governed countries, I discount the possibility that high biomass in these distant, productive regions could be the result of better modern governance. The same countries which target Alaska and the European Arctic, namely the US and European countries, also target areas which are more exploited and do not have high biomass or high rates of MSC enrollment, such as the US East Coast and the Mediterranean. I therefore conclude that a typical, high-propensity MSC fishery is likely to already be well-regulated and target an already-healthy stock.

While the enrollment model was dominated by these sort of cost variables, i.e. how close a fishery already was to MSC standards, the withdrawal model was determined by variables which suggested greater benefits. The most important factors in the enrollment model, after regional harvest, were size, price, and trade. These last two, I interpret, reflect how export-oriented a fishery is (Gollin, 2014). Highly-traded fish species landed by countries which have high seafood exports and export-heavy trade balances are much more likely to end up on the international market. Similarly, high-priced seafood, no matter where it is caught, tends to be sold to and consumed by the Global North (Asche et al., 2015). On the globally-competitive international market, MSC-certified fisheries may be more likely to fetch a premium, greater market share, or other benefit, by having access to more rich-world consumers (Bellchambers et al., 2016; Oosterveer, 2008). I therefore contend that export-oriented fisheries, particularly in the Global South, are more likely to realize benefits from MSC certification, and therefore to remain in the program.

I attribute the disparity between the enrollment and withdrawal decisions to asymmetric information available to fishery decision-makers. Given the uncertainty around price premiums,even in academic circles (Bellchambers et al., 2016; Roheim et al., 2011; Sánchez et al., 2020; Wakamatsu, 2014), decision-makers are unlikely to know in advance what benefits they are likely to receive from an MSC certification. As a result, MSC’s enrollment seems to be dominated by fisheries with low up-front costs, as those which require fewer changes (Martin et al., 2012) are likely more willing to take a chance and enroll in a program whose benefits are unclear or unproven. This lack of information skews enrollment towards already-sustainable and well-managed fisheries, and may also explain why attrition rates are so high. Many fisheries enter the program without clear knowledge of what benefits they will receive, and only those for whom benefits materialize remain. Asymmetric information, therefore, may be a major factor in MSC’s current inability to reach those fisheries most in need of improvements.

Addressing this lack of information around expected benefits could have strong implications for the potential of certification programs to contribute to sustainability in contexts where they may be most needed. While MSC’s enrollment is dominated by well-regulated fisheries in sustainable regions, atypical MSCs which do not follow this mould are not necessarily less likely to remain in and realize benefits from the program. As a result, there may be a large number of fisheries which could potentially be viable and successful MSC programs, but are discouraged by high up-front costs without knowledge of the benefits they stand to receive. I therefore contend that the MSC may face fewer systemic obstacles in efforts to broaden their enrollment than is often assumed. If the MSC can improve the availability of information around expected benefits, and identify export-oriented fisheries in the Global South for whom these benefits exceed the potentially-daunting initial costs (Pérez-Ramírez et al., 2012a), they may succeed in creating a more diverse enrollment of fisheries, which, crucially, are also in greater need of improvements.

Finally, the outcomes model identified desirable stock assessment outcomes for both catch and biomass in MSC fisheries, relative to certain controls, although available data limited my ability to examine a wide range of MSC certifications. Biomass improvements were greatest around the time of certification, but were evident throughout the first five years of certification, relative to both high-propensity and all untreated controls. We attribute the initial peak to the effects of assessment. Because fisheries pay in advance for the certification process without any guarantees as to the outcome, they may go above and beyond in terms of improvements or behavior in order to ensure a positive outcome from their investment. Improvements during this 'best behavior' period and in the medium-term four years after certification did not correlate with enrollment or withdrawal propensity, but data were only available for a fairly 'typical' subset of MSCs (Table 3). I relied exclusively on publicly-available stock assessments, and could only use those which had data beginning at least 8 years before certification. This limits my findings to a particular subset of MSC fisheries, as stocks with longstanding public assessments are likely already well-understood and receive significant attention from regulators, due to the expense of collecting this kind of data. However, even within these fisheries, we still detected significant improvements as a result of MSC certification, likely reflecting MSC rules acting as a supplement, rather than an alternative, to conventional regulation, consistent with Hønneland et al. (2020).

While catch figures did not decrease in absolute terms in enrolled fisheries, total landings remained, on average, constant to slightly declining, while those of unenrolled synthetic control fisheries increased dramatically. However, catch in high-propensity control stocks, those unenrolled fisheries which have similar socio-economic contexts as MSCs, also remained relatively constant, so it is unclear if restrained catches in MSC fisheries are due to the effects of the program or strong background regulation. As catch did not consistently decrease, the observed increases in MSC biomass over the same period, which also held when compared with high-propensity controls, is likely due therefore to changes in fishing practices, such as gear use, bycatch reduction, targeting different areas etc., rather than absolute reduction in landings. These stock assessment patterns may indicate socially-desirable outcomes for MSC-enrolled fisheries, with fishers able to realize higher prices on similar landings while allowing biomass to increase. However, greater investigation is needed and data limitations must be overcome before these patterns can be directly attributed to MSC enrollment. It is unclear if MSC fisheries in low-propensity contexts would be able to restrain catch and lead to similar biomass improvements as it did in our higher-propensity subset, given the differences in catch data between unenrolled fisheries in high-propensity vs. all other contexts.

References:

Abadie, A., Diamond, A., Hainmueller, J., 2010. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. J. Am. Stat. Assoc. 105, 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Allison, E.H., Ellis, F., 2001. The livelihoods approach and management of small-scale fisheries. Mar. Policy 25, 377–388. https://doi.org/10.1016/S0308-597X(01)00023-9

Andrew, N.L., Béné, C., Hall, S.J., Allison, E.H., Heck, S., Ratner, B.D., 2007. Diagnosis and management of small-scale fisheries in developing countries. Fish Fish. 8, 227–240. https://doi.org/10.1111/j.1467-2679.2007.00252.x

Asche, F., Bellemare, M.F., Roheim, C., Smith, M.D., Tveteras, S., 2015. Fair Enough? Food Security and the International Trade of Seafood. World Dev. 67, 151–160. https://doi.org/10.1016/j.worlddev.2014.10.013

Bellchambers, L.M., Phillips, B.F., Pérez-Ramírez, M., 2016. From certification to recertification the benefits and challenges of the Marine Stewardship Council (MSC): a case study using lobsters. Fish. Res. 182, 88–97.

Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. Classification and regression trees. CRC press.

Christian, C., Ainley, D., Bailey, M., Dayton, P., Hocevar, J., LeVine, M., Nikoloyuk, J., Nouvian, C., Velarde, E., Werner, R., Jacquet, J., 2013. A review of formal objections to Marine Stewardship Council fisheries certifications. Biol. Conserv. 161, 10–17. https://doi.org/10.1016/j.biocon.2013.01.002

Coglianese, C., Nash, J., 2001. Regulating from the inside: can environmental management systems achieve policy goals? Resources for the Future.

Cole, M.A., Elliott, R.J., Liu, B., 2020. The impact of the Wuhan Covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach. Environ. Resour. Econ. 76, 553–580.

Couronné, R., Probst, P., Boulesteix, A.-L., 2018. Random forest versus logistic regression: a large-scale benchmark experiment. BMC Bioinformatics 19, 270. https://doi.org/10.1186/s12859-018-2264-5

FAO, 2018. The State of World Fisheries and Aquaculture 2018-Meeting the sustainable development goals. Licence CC -NC-SA 30 IGO.

Fisheries - MSC Fisheries [WWW Document], n.d. URL https://fisheries.msc.org/en/fisheries/ (accessed 6.9.21).

FishSource [WWW Document], n.d. URL https://www.fishsource.org/ (accessed 6.9.21).

Foley, P., 2012. The Political Economy of Marine Stewardship Council Certification: Processors and Access in Newfoundland and Labrador's Inshore Shrimp Industry. J. Agrar. Change 12, 436–457. https://doi.org/10.1111/j.1471-0366.2011.00344.x

Fung, A., Graham, M., Weil, D., Fagotto, E., 2007. Transparency policies: Two possible futures. Taubman Cent. Policy Briefs 1–6.

Gollin, D., 2014. The Lewis Model: A 60-Year Retrospective. J. Econ. Perspect. 28, 71–88. https://doi.org/10.1257/jep.28.3.71

Goyert, W., Sagarin, R., Annala, J., 2010. The promise and pitfalls of Marine Stewardship Council certification: Maine lobster as a case study. Mar. Policy 34, 1103–1109. https://doi.org/10.1016/j.marpol.2010.03.010

Gulbrandsen, L.H., 2009. The emergence and effectiveness of the Marine Stewardship Council. Mar. Policy 33, 654–660. https://doi.org/10.1016/j.marpol.2009.01.002

Gutiérrez, N.L., Valencia, S.R., Branch, T.A., Agnew, D.J., Baum, J.K., Bianchi, P.L., Cornejo-Donoso, J., Costello, C., Defeo, O., Essington, T.E., Hilborn, R., Hoggarth, D.D., Larsen, A.E., Ninnes, C., Sainsbury, K., Selden, R.L., Sistla, S., Smith, A.D.M., Stern-Pirlot, A., Teck, S.J., Thorson, J.T., Williams, N.E., 2012. Eco-Label Conveys

Reliable Information on Fish Stock Health to Seafood Consumers. PLoS ONE 7, e43765. https://doi.org/10.1371/journal.pone.0043765
Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.
Hilborn, R., Amoroso, R.O., Anderson, C.M., Baum, J.K., Branch, T.A., Costello, C., Moor, C.L. de, Faraj, A., Hively, D., Jensen, O.P., Kurota, H., Little, L.R., Mace, P., McClanahan, T., Melnychuk, M.C., Minto, C., Osio, G.C., Parma, A.M., Pons, M., Segurado, S., Szuwalski, C.S., Wilson, J.R., Ye, Y., 2020. Effective fisheries management instrumental in improving fish stock status. Proc. Natl. Acad. Sci. 117, 2218–2224. https://doi.org/10.1073/pnas.1909726116
Hønneland, G., 2020. Marine Stewardship Council (MSC) Certification of Arctic Fisheries: Processes and Outcomes. Arct. Rev. 11, 133–156. https://doi.org/10.23865/arctic.v11.2488
Izquierdo-Peña, V., Lluch‐Cota, S.E., Chavez, F.P., Lluch‐Cota, D.B., Morales-Bojórquez, E., Ponce-Díaz, G., 2020. Is there a Future in the Sustainability Certification of Sardine and Anchovy Fisheries? Fisheries 45, 554–560. https://doi.org/10.1002/fsh.10461
Jardine, S.L., Sanchirico, J.N., 2012. Catch share programs in developing countries: A survey of the literature. Mar. Policy 36, 1242–1254.
Kalfagianni, A., Pattberg, P., 2013. Global fisheries governance beyond the State: unraveling the effectiveness of the Marine Stewardship Council. J. Environ. Stud. Sci. 3, 184–193. https://doi.org/10.1007/s13412-013-0118-z
King, G., Nielsen, R., 2019. Why Propensity Scores Should Not Be Used for Matching. Polit. Anal. 27, 435–454. https://doi.org/10.1017/pan.2019.11
Kuosmanen, T., Zhou, X., Eskelinen, J., Malo, P., 2021. Design Flaw of the Synthetic Control Method.
Manach, F.L., Jacquet, J.L., Bailey, M., Jouanneau, C., Nouvian, C., 2020. Small is beautiful, but large is certified: A comparison between fisheries the Marine Stewardship Council (MSC) features in its promotional materials and MSC-certified fisheries. PLOS ONE 15, e0231073. https://doi.org/10.1371/journal.pone.0231073
Martin, S.M., Cambridge, T.A., Grieve, C., Nimmo, F.M., Agnew, D.J., 2012. An Evaluation of Environmental Changes Within Fisheries Involved in the Marine Stewardship Council Certification Scheme. Rev. Fish. Sci. 20, 61–69. https://doi.org/10.1080/10641262.2011.654287
Oosterveer, P., 2008. Governing global fish provisioning: Ownership and management of marine resources. Ocean Coast. Manag. 51, 797–805. https://doi.org/10.1016/j.ocecoaman.2008.08.002
Pascoe, S., Innes, J., Holland, D., Fina, M., Thébaud, O., Townsend, R., Sanchirico, J., Arnason, R., Wilcox, C., Hutton, T., 2010. Use of incentive-based management systems to limit bycatch and discarding. Int. Rev. Environ. Resour. Econ. 4, 123–161.
Pérez-Ramírez, M., Castrejón, M., Gutiérrez, N.L., Defeo, O., 2016. The Marine Stewardship Council certification in Latin America and the Caribbean: A review of experiences, potentials and pitfalls. Fish. Res. 182, 50–58.
Pérez-Ramírez, M., Phillips, B., Lluch-Belda, D., Lluch-Cota, S., 2012a. Perspectives for implementing fisheries certification in developing countries. Mar. Policy 36, 297–302.
Pérez-Ramírez, M., Ponce-Díaz, G., Lluch-Cota, S., 2012b. The role of MSC certification in the empowerment of fishing cooperatives in Mexico: The case of red rock lobster co-managed fishery. Ocean Coast. Manag. 63, 24–29.
Ponte, S., 2012. The Marine Stewardship Council (MSC) and the Making of a Market for ‘Sustainable Fish.’ J. Agrar. Change 12, 300–315. https://doi.org/10.1111/j.1471-0366.2011.00345.x
Prakash, A., Potoski, M., 2007. Collective action through voluntary environmental programs: A

club theory perspective. Policy Stud. J. 35, 773–792.
Roheim, C.A., Asche, F., Santos, J.I., 2011. The Elusive Price Premium for Ecolabelled Products: Evidence from Seafood in the UK Market. J. Agric. Econ. 62, 655–668. https://doi.org/10.1111/j.1477-9552.2011.00299.x
Roheim, C.A., Sutinen, J.G., 2006. Trade and Marketplace Measures to Promote Sustainable Fishing Practices. https://doi.org/10.7215/nr_ip_20060501
Sampson, G.S., Sanchirico, J.N., Roheim, C.A., Bush, S.R., Taylor, J.E., Allison, E.H., Anderson, J.L., Ban, N.C., Fujita, R., Jupiter, S., Wilson, J.R., 2015. Secure sustainable seafood from developing countries. Science 348, 504–506. https://doi.org/10.1126/science.aaa4639
Sánchez, J.L.F., Polanco, J.M.F., García, I.L., 2020. Evidence of price premium for MSC-certified products at fishers' level: The case of the artisanal fleet of common octopus from Asturias (Spain). Mar. Policy 119, 104098.
Stratoudakis, Y., McConney, P., Duncan, J., Ghofar, A., Gitonga, N., Mohamed, K.S., Samoilys, M., Symington, K., Bourillon, L., 2016. Fisheries certification in the developing world: Locks and keys or square pegs in round holes? Fish. Res., Special Issue: Fisheries certification and Eco-labeling: Benefits, Challenges and Solutions 182, 39–49. https://doi.org/10.1016/j.fishres.2015.08.021
Thomas Travaille, K.L., Lindley, J., Kendrick, G.A., Crowder, L.B., Clifton, J., 2019. The market for sustainable seafood drives transformative change in fishery social-ecological systems. Glob. Environ. Change 57, 101919. https://doi.org/10.1016/j.gloenvcha.2019.05.003
Uchida, H., Roheim, K., Bush, S., Sanchirico, J., Asche, F., Zou, C., 2017. A New Theory of Change for the Sustainable Seafood Movement.
Varian, H.R., 2014. Big Data: New Tricks for Econometrics. J. Econ. Perspect. 28, 3–28. https://doi.org/10.1257/jep.28.2.3
Wakamatsu, H., 2014. The Impact of MSC Certification on a Japanese Certified Fishery. Mar. Resour. Econ. 29, 55–67. https://doi.org/10.1086/676287
Xu, Y., 2017. Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. Polit. Anal. 25, 57–76. https://doi.org/10.1017/pan.2016.2

Figure Index: