NBER WORKING PAPER SERIES

NON-PARAMETRIC TESTS OF THE TRAGEDY OF THE COMMONS

H. Spencer Banzhaf
Yaqin Liu
Martin Smith
Frank Asche

Working Paper 26398
http://www.nber.org/papers/w26398

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2019

We thank Anna Birkenbach, Rahul Deb, Matt Kotchen, Chuck Mason, Kurt Schnier, Jeff Wooldridge and participants at Camp Resources XX, the 2018 Southern Economic Association conference, the 2018 World Congress of Environmental and Resource Economists, and the 2019 AERE Conference for helpful comments and discussions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w26398.ack

Non-Parametric Tests of the Tragedy of the Commons
H. Spencer Banzhaf, Yaqin Liu, Martin Smith, and Frank Asche
NBER Working Paper No. 26398
October 2019
JEL No. C14,C72,D21,D23,Q2,Q3

ABSTRACT

Extending recent results in the industrial organization literature (Carvajal et al. 2013), we de-rive non-parametric tests of behavior consistent with the tragedy of the commons model. Our approach derives testable implications of such behavior under any arbitrarily concave, differentiable production function of total inputs and when individual extractors of the resource have any arbitrary convex, differentiable cost of supplying inputs. We extend the tests to account for behavioral errors in observed data and derive statistical tests based on "how far off" the marginal costs are from those that are consistent with the model. We also extend the tests to allow for sampling error and/or measurement error. Applying our approach to panel data of Norwegian fishers, we find evidence rejecting the tragedy of the commons model. Significantly, we find that rejection rates of the model increase after property rights reforms moved the fishery away from the tragedy of the commons.

H. Spencer Banzhaf
Department of Economics
Andrew Young School of Policy Studies
Georgia State University
P.O. Box 3992
Atlanta, GA 30302
and NBER
hsbanzhaf@gsu.edu

Yaqin Liu
Nicolas School of the Environment
Duke University
Durham, NC 27708
yliu64@student.gsu.edu

Martin Smith
Nicholas School
Duke University
Box 90328
Durham, NC 27708
marsmith@duke.edu

Frank Asche
University of Florida
frank.asche@ufl.edu

Non-Parametric Tests of the Tragedy of the Commons

Spencer Banzhaf, Yaqin Liu, Martin Smith, and Frank Asche

1. Introduction

The "tragedy of the commons" (Hardin 1968) occurs when strategic incentives, unchecked by property rights or other institutional arrangements, undermine the potential value of a commonly held resource. Because individuals do not bear the full cost when they utilize the common resource, they have an incentive to use it too intensively, relative to the group's welfare. In the standard model, individuals receive a prorated share of collective output, proportionate to their inputs, so by increasing inputs they can obtain a larger share of the pie (Gordon 1954, Weitzman 1974, Dasgupta and Heal 1979). Classic examples include sending cattle to a common pasture (Huffaker and Wilen 1991), cooperatives (Sen 1966), extracting oil from a common pool (Libecap and Wiggins 1984, Baltrop and Schnier 2016), extracting groundwater (Brazović et al. 2010, Koch and Nax 2017, Burlig et al. 2019, Ayres et al. 2019), and fishing from the sea (Gordon 1954; Costello et al. 2008, Huang and Smith 2014, Birkenbach et al. 2017). Stavins (2011) offers a review.

Though examples of the tragedy at work are pervasive, groups can avoid the trap of open-access by devising ways to cooperate and limit access to the commons, effectively managing common-pool resources to avoid the tragedy (Ciriacy-Wantrup and Bishop 1975, Ostrom 1990). Evidence from laboratory experiments suggests that when they make decisions anonymously and without communication, individuals do over-exploit common resources, producing the "tragedy," but when they can communicate and/or can build other institutions to change incentives, they can overcome the tragedy (Ostrom 2009).

Surprisingly, then, there have been few empirical tests of the standard model with naturally occurring data. Several studies have considered the aggregate effects of different property rights regimes. In the context of pumping races, Balthrop and Schnier (2016) find that unitization decreases the race to pump. In the context of fisheries, Costello et al. (2008) and Birkenbach et al. (2017) find that individual catch shares can prevent the collapse of fisheries and slow the race to fish. These policy outcomes are consistent with over-exploitation in the open-access regime, but do not test individual behavior.

Kirkley et al. (2002) and Felthoven et al. (2009) outline approaches for measuring capacity utilization in an industry exploiting a common pool resource, such as a fishery, interpreting excess capacity as a symptom of the excessive application of variable inputs to the resource. This approach requires estimating a production function for firms. But, although they certainly estimate important policy effects of various property rights regimes, and although they provide "circumstantial" evidence of commons-like behavior, none of these papers provide an explicit mapping from the strategic behavior in the commons model to the data in a way which allows the behavioral model to be tested.

Taking a very different approach, Huang and Smith (2014) conducted the first micro-level empirical investigation of strategic behavior in a common pool. They develop a dynamic structural model of the microeconomic behavior of fishers operating in an open access fishery. Each fisher chooses his effort to maximize his expected utility given all other fishers' actions, with agglomeration or congestion effects specified such that individual catch per day is affected by the total number of vessels fishing on that day. With estimates from their parametric model, potential efficiency gains can be quantified by comparing the optimal vessel numbers to the predicted numbers resulting from the individual maximization problem. However, their approach presupposes Nash behavior in a commons game rather than providing a way to test for such behavior. Moreover, their approach is highly parametric, which has the advantage of allowing for counter-factual policy simulations and welfare analyses, but comes at the cost of bringing in numerous maintained assumptions when it comes to testing for particular modes of strategic behavior.

In this paper, we introduce an alternative empirical strategy that complements the existing literature. In particular, we develop a non-parametric revealed preference-type test for the canonical behavioral model of the tragedy of the commons. Recently, Carvajal et al. (2013) developed a revealed preference test for Cournot equilibrium, deriving properties that hold when firms are strategically interacting as predicted by that model. As the tragedy of the commons and the Cournot model are essentially isomorphic (both are surplus-sharing games), we derive similar properties that hold under the strategic interactions of the tragedy of the commons. Our test has the advantage of requiring no parametric assumptions about production functions or cost functions (beyond convexity). The test is derived from the key characteristics of the tragedy of the commons that each agent maximizes its objective function independently and from a proportionate sharing rule. The test can be implemented with panel data of individual inputs and total output. In particular, given panel data on each agent's input and the total output from exploitation, we show that a data set is consistent to the tragedy of the commons with convex cost functions if and only if there is a solution to a linear program that we can explicitly construct from the data. Accordingly, the tests we derive can be applied to various settings with common pool resources, from fisheries to oil and water extraction.

Beyond adapting the approach of Carvajal et al. to the commons, we extend their tests to incorporate sampling errors in total input and output. Sampling error is modeled as a latent parameter, which can be inferred from our linear program under the null hypothesis of behavior consistent with the tragedy-of-the-commons. The model allows for the analyst to impose boundaries on permissible sampling errors based on credible information or assumptions. Sampling errors change the testable properties, and increase the domain of the linear program, which make the test less stringent. Hence, compared to the basic tests, tests with sampling errors reduce rejection rates of the model.

Additionally, we derive tests to gauge the minimum distance of the set of recovered marginal costs from those that are consistent with the model. Developing ideas proposed by Afriat (1972), Diewert (1973), and Varian (1985), we include an adjustment factor in the model to guarantee that data would always pass the behavioral test. We apply a linear program to reveal the minimal magnitude of the adjustments required as a measure of distance from the model. In one version of this approach, we consider behavioral errors in which the marginal costs used in the firms' objective functions depart from the true costs. In another version, we consider measurement error in inputs. Using these errors, we apply a Kolmogorov-Smirnov test to inform probability distributions for rejections of the model. These extensions also could be applied to the tests of the Cournot model (as in Carvajal et al.) as well as the tragedy of the commons.

We take the test to the Norwegian coastal fishery for cod and other whitefish (the largest fishery in Norway and a major contributor to the global market for whitefish). Our basic results reject behavior consistent with the tragedy of the commons using the full data sets. Results from tests with sampling errors display lower rejection rates in general but do not alter the pattern. Significantly, preliminary results show that the rejection rates are higher after property-rights reforms in the Norwegian fishery that reduced open-access incentives. In other words, using our test, the tragedy of the commons model is rejected to a greater degree after these reforms, as we would expect.

The rest of the paper is organized as follows. In Section 2, we derive the theoretical results for the classic static model of the average return game, in which agents select their inputs and each unit of input receives the average return (rather than marginal return). In Section 3, we offer additional extensions to the model, including quantifying distance to the model, conducting statistical tests, and measurement error. Section 4 discusses the empirical application and Section 5 shows the results. Section 6 concludes.

2. Basic Result: A Nonparametric Test of the Tragedy of the Commons

2.1. The Static Average Return Game

Consider an industry consisting of I profit-maximizing firms, indexed by $i = 1,2,\dots,I$, each having free access to an exogenously fixed common property resource. There are T decision periods indexed by $t = 1,2,\dots,T$. Denote $q_{i,t}$ as the extraction effort by firm i in period t. For example, $q_{i,t}$ might be the number of fishing vessel-days in year t. Let $Q_t = \sum_i q_{i,t}$ be the total level of effort applied to the resource at time t. The differentiable production function for the industry at time t is $Y_t = F_t(Q_t)$, with $F(0) = 0$, $F'(Q) > 0$, and F' non-increasing for all t.

Following the canonical commons model (Gordon 1954, Weitzman 1974, Dasgupta and Heal 1979, Cornes and Sandler 1996), each firm's extraction is proportionate to its share of input. Thus, firm i's revenue in period t is $\frac{q_{i,t}}{Q_t} * p_t F_t(Q_t)$, where p_t denotes the market price of output (e.g. fish) at time t. This assumption captures the characteristic of open-access resources that factors tend to receive their average rather than the marginal product. Finally, let $C_i(q_{i,t})$ denote firm i's cost function, which is a differentiable and non-decreasing function of q.

Following Carvajal et al.'s logic for Cournot competition, we say a panel data set $\mathcal{O} = \left\{p_t F_t, \left(q_{i,t}\right)_{i\epsilon 1\dots N}\right\}_{t\epsilon 1\dots T}$ is consistent with the tragedy of the commons if there exist cost functions $\bar{C}_i$ for each firm i, and concave production functions $\bar{F}_t$ for each observation t which jointly satisfy the following two conditions:

(i) $\bar{F}_t(Q_t) = F_t$

(ii) $q_{i,t} \epsilon\ argmax_{\tilde{q}_{i,t} \geq 0}\{\frac{\tilde{q}_{i,t}}{Q_t} * p_t F_t(Q_t) - \bar{C}_i(\tilde{q}_{i,t})\}.$

Condition (i) says the production function must be consistent with observed output at time *t*. Condition (ii) says firm *i*'s input at time *t* maximizes its profit given the inputs of all other firms (a standard Nash assumption).

Note that we do not need to estimate the production function. We allow the analysis to explain the data using any arbitrary concave production function, as long as it passes through the observed total output and inputs, $p_t F_t(Q_t)$ and Q_t, at each decision period. Similarly, no restrictions are placed on firms' cost functions except that they are increasing and convex.

To see how we can avoid functional form assumptions, consider firm *i*'s profit-maximization problem at time t:

$$\max_{q_{i,t}} \frac{q_{i,t}}{Q_t} * p_t F_t(Q_t) - C_i\,(q_{i,t}). \quad (1)$$

Taking other firms' actions as given, the first-order condition is:

$$\frac{q_{i,t}}{Q_t} * p_t F_t'(Q_t) + \left(1 - \frac{q_{i,t}}{Q_t}\right) * \frac{p_t F_t(Q_t)}{Q_t} = C_{i,t}'. \quad (2)$$

This is the standard result that firms equate marginal cost to a weighted average of marginal returns and average returns (Weitzman 1974, Dasgupta and Heal 1979). In the case of a monopolist, $q_{i,t} = Q_t$ and the entire weight is on the efficient condition to equate marginal cost to marginal return. In the limit, as the firms grows small, $q_{i,t}/Q_t$ goes to zero and the firms equate marginal cost to average revenue, thus depleting all resource rents (as in Gordon 1954).

Rearranging terms, we obtain:

$$\frac{p_t F_t(Q_t) - Q_t C_{i,t}'}{q_{i,t}} = \frac{p_t F_t(Q_t)}{Q_t} - p_t F_t'(Q_t). \quad (3)$$

Notice in Equation (3) that the left-hand side involves firm-specific terms (inputs $q_{i,t}$ and marginal costs $C_{i,t}'$) while the right-hand side involves only market-wide data (total revenue $p_t F_t(Q_t)$, marginal revenue product $p_t F_t'$, and total input Q_t). Consequently, from the first-order condition, we obtain a *common ratio property* comparable to that in Carvajal et al.:

(4) $\frac{p_t \mathrm{F}_t(Q_t) - Q_t C'_{i,t}}{q_{i,t}} = \frac{p_t \mathrm{F}_t(Q_t) - Q_t C'_{j,t}}{q_{j,t}} = \cdots = \frac{p_t \mathrm{F}_t(Q_t) - Q_t C'_{I,t}}{q_{I,t}} \geq 0$ for $t \in T$.

In other words, in each period, functions of firms' extraction effort and marginal costs should all be equal. The expressions are nonnegative given the concavity of the production function.

Moreover, because each firm's cost function is convex, the array $\{C'_{i,t}\}$ displays increasing marginal costs for each firm *i*. Thus, if the cost function is time-invariant, we also have the *co-monotone property* as described in Carvajal et al., such that for all *i*,

(5) $\quad q_{i,t} > q_{i,t'} \rightarrow C'_{i,t} \geq C'_{i,t'}$.

Consequently, a set of observations is consistent with the tragedy of the commons with convex cost functions if and only if there exist nonnegative numbers $\{C'_{i,t}\}$ for all *i*,*t* that obey the common ratio and co-monotone properties. In Example 1, we show that certain data sets are not consistent with the tragedy of the commons given the interplay of the two properties.

Example 1: Consider the following observations of two firms i and j sharing a common-pool resource:

(i) At observation t, $p_t F_t(Q_t) = 50$, $q_{i,t} = 50$, $q_{j,t} = 100$.

(ii) At observation t', $p_{t'} F_{t'}(Q_{t'}) = 350$, $q_{i,t'} = 70$, $q_{j,t'} = 60$.

Re-arranging the common-ratio property at *t'* to isolate $C'_{j,t'}$ and using the fact that $\frac{q_{j,t'}}{q_{i,t'}} C'_{i,t'} \geq 0$, we have:

$$C'_{j,t'} = \frac{p_{t'} F_{t'}(Q_{t'})}{Q_{t'}} - \frac{q_{j,t'}}{q_{i,t'}} \frac{p_{t'} F_{t'}(Q_{t'})}{Q_{t'}} + \frac{q_{j,t'}}{q_{i,t'}} C'_{i,t'} \geq \frac{p_{t'} F_{t'}(Q_{t'})}{Q_{t'}} - \frac{q_{j,t'}}{q_{i,t'}} \frac{p_{t'} F_{t'}(Q_{t'})}{Q_{t'}} = 0.385.$$

Now, we know from the first-order condition (2) that $C'_{i,t} < \frac{p_t F_t(Q_t)}{Q_t}$, at each time *t* for all *i*, because $C'_{i,t} = \frac{q_{i,t}}{Q_t}\left(p_t F'_t(Q_t) - \frac{p_t F_t(Q_t)}{Q_t}\right) + \frac{p_t F_t(Q_t)}{Q_t}$ and $F'_t(Q_t) - \frac{F_t(Q_t)}{Q_t} < 0$ given the concavity of production function. Thus, $C'_{j,t} < \frac{p_t F_t(Q_t)}{Q_t} = 0.33$. In addition, from the co-monotone property, we have $C'_{j,t'} \leq C'_{j,t}$ because $q_{j,t'} < q_{j,t}$. Thus, in sum, $0.385 \leq C'_{j,t'} < C'_{j,t} < 0.33$, which is clearly a contradiction. Thus, there are no nonnegative marginal costs that satisfy the common-ratio property and the co-monotone properties. The data in Example 1 are not consistent with the tragedy of the commons model.

2.2. Implementation: A Linear Program for the Test

Our approach to testing the tragedy-of-the-commons model can be reformulated as a simple linear program: Given panel data on each agent's input and the total output from exploitation, find nonnegative marginal costs, $\{C'_{i,t}\}$, for all agents *i* at each time *t*, which satisfy the common-ratio property (4) and the co-monotone property (5). This linear program is analogous to the conditions specified in Afriat's Theorem for testing whether consumers' choices are consistent with utility-maximizing behavior or, equivalently, the Generalized Axiom of Revealed Preference (GARP) (Afriat 1967). This overall approach encompasses a diversity of research programs and has been extended to a wide array of settings (Chambers and Echenique 2016, Hands 2014), including firms' costs (Varian 1984) and Cournot competition (Carvajal et al. 2013).

In our context, a set of observations is consistent with the tragedy of the commons with convex cost functions if and only if, given the observed $p_t \mathrm{F}_t$, $q_{i,t}$, and Q_t there are numbers $C'_{i,t}$ satisfying:

(i) $\frac{p_t \mathrm{F}_t(Q_t) - Q_t C'_{i,t}}{q_{i,t}} = \frac{p_t \mathrm{F}_t(Q_t) - Q_t C'_{j,t}}{q_{j,t}} \geq 0 \;\; \forall\, i,j \in I, \forall\, t \in T;$

(ii) $(q_{i,t} - q_{i,t'})(C'_{i,t} - C'_{i,t'}) \geq 0 \;\; \forall\, i \in I, \forall\, t, t' \in T;$

(iii) $C'_{i,t} \geq 0 \;\; \forall\, i \in I, \forall\, t \in T.$

See Appendix A for a proof.

Condition (i) is the common-ratio property which follows from the first-order condition; condition (ii) is the co-monotone property which follows from the convexity of the cost function; and condition (iii) is a non-negativity constraint which follows from the fact that the cost function is increasing. For a panel data set, failure to obtain a solution to any element in the marginal cost set $\{C'_{i,t}\}_{\forall\, i \in I, \forall\, t \in T}$, will result in a rejection of the model.

To understand the implications of this test, we emphasize three features. First, it is entire *data sets* that are or are not rejected, not individual observations or individual firms. Again, this feature is consistent with tests of consumers' choices, in which entire data sets are or are not consistent with GARP, not individual choices. However, one can always throw out particular observations from the data set and consider the effect of doing so. Thus, taking random subsets of the data, one can generate rejection *rates*, as a quantitative measure of "how much" the data are inconsistent with the tragedy of the commons model. Further, one can isolate data from particular firms or periods to see if the data set is more likely to be rejected with or without them. Below, we leverage this possibility in our empirical applications to test the effect on rejection rates of including data generated under differing property rights regimes.

Second, our approach tests the minimum *necessary* conditions for the above behavioral model. Under the model's behavioral assumptions, the test eliminates *any* type I error. On the other hand, it is weak in the sense of potentially allowing a great deal of type II error. That is, rejection of the model gives one confidence that the data indeed are not consistent with the tragedy of the commons model, but—as always—failure to reject does not guarantee the model is true (nor, of course, that alternative models are false). This is not a limitation of our approach so much as a limitation of what can be said about the behavioral model: if further restrictions would lead to more rejections, then arguably it is the auxiliary hypotheses that are being rejected, not the fundamentals of the behavioral model. It is always the case that failure to reject a null hypothesis does not guarantee it to be true.

Third, nevertheless, even with the very weak assumptions we bring to the model, we still can learn a great deal from the tests derived from it. Data sets that are consistent with the tragedy of the commons model are inconsistent with at least some rival models. Consider, for example, the case of non-tradable quotas, which restrict each firm to extract only up to its quota. Although non-tradability prevents cost minimization subject to total extraction by the group (as firms with high costs at the margin may be allocated quota that cannot be traded to low-cost firms), non-tradable quotas do have some advantages. Typically, they cap the total allowable extraction so as to protect the sustainability of a resource. Additionally, unlike group quotas (which also cap total extraction), they prevent a "race" within the time period over which the quota is defined, as a firm's share is exogenous to how quickly it extracts. This can prevent, e.g., a race to pump water or oil or to catch fish in order to get a larger share of the group quota.

Importantly, non-tradable quotas do not lead to a common ratio property like Equation (4). To see this, note that the objective function would now be written as a constrained optimization problem:

$$\text{(1')} \qquad \max_{q_{i,t}} \frac{q_{i,t}}{Q_t} * p_t F_t(Q_t) - C_i\left(q_{i,t}\right) + \lambda_{i,t}\left(L_{i,t} - \frac{q_{i,t}}{Q_t} * F_t(Q_t)\right),$$

where $L_{i,t}$ is the quota limit and $\lambda_{i,t}$ is the shadow cost of that limit. Note output prices appear in the revenue term but not the constraint. The revised first-order condition is:

$$\text{(2')} \qquad \left(p_t - \lambda_{i,t}\right)\left[\frac{q_{i,t}}{Q_t} * F_t'(Q_t) + \left(1 - \frac{q_{i,t}}{Q_t}\right) * \frac{F_t(Q_t)}{Q_t}\right] = C_{i,t}'.$$

The quota is associated with a firm-specific shadow price on extraction, so it is equivalent to the original problem with an adjusted output price. Finally, rearranging terms, we obtain:

$$\text{(3')} \qquad \frac{F_t(Q_t) - Q_t C_{i,t}'/(p_t - \lambda_{i,t})}{q_{i,t}} = \frac{F_t(Q_t)}{Q_t} - F_t'(Q_t).$$

Taking this equation in isolation, it would appear that instead of solving the linear program by finding numbers $C_{i,t}'$, we could instead simply solve for numbers $C_{i,t}'/\left(p_t - \lambda_{i,t}\right)$. However, the latter numbers would not be expected to satisfy the co-monotone property, which is based on the convexity of $C_{i,t}'$ alone. For example, ceteris paribus, higher effort one year might come with a higher quota, but this would tend to lower $\lambda_{i,t}$ (as the quota is less binding), and hence lower the over-all expression $C_{i,t}'/\left(p_t - \lambda_{i,t}\right)$, perhaps violating the co-monotone property.

Thus, we would expect an IVQ regime to lead to higher rejection rates. We leverage this insight in our empirical work below.

3. Extensions

In this section, we extend the model in various ways. Our extensions can be applied to other settings as well, including the case of Cournot competition considered by Carvajal et al. (2013). Thus, they represent an additional contribution of this research.

3.1 The Test with Sampling Error

The test we derived in Section 2 assumes that data are observed without error. Moreover, it assumes data from a *census* (not just sample) of users, so that $Q = \sum_i q_i$ and total catch $F(Q)$ are observed. In this section, we consider the case where only a sample of users are observed, so that total effort Q and total revenue F are estimates based on sample mean times N.

If total effort and total revenue are based on sample averages, they are observed with error. Let α_t and β_t be the respective proportionate errors, so we observe $p_t\hat{F}_t = p_t F_t * \alpha_t$ and $\hat{Q}_t = Q_t * \beta_t$. Then the common ratio property becomes $\frac{\alpha_t p_t \mathrm{F}_t(Q_t) - \beta_t Q_t(C'_{i,t})}{q_{i,t}} = \frac{\alpha_t p_t \mathrm{F}_t(Q_t) - \beta_t Q_t(C'_{j,t})}{q_{j,t}}$. Dividing both sides by β_t and letting $\gamma_t = \alpha_t/\beta_t$, we can write the linear program with sampling errors as:

(i) $\frac{\gamma_t p_t \mathrm{F}_t(Q_t) - Q_t(C'_{i,t})}{q_{i,t}} = \frac{\gamma_t p_t \mathrm{F}_t(Q_t) - Q_t(C'_{j,t})}{q_{j,t}} \geq 0,\ \forall\, i,j \in I, \forall\, t \in T;$

(ii) $(q_{i,t} - q_{i,t'})(C'_{i,t} - C'_{i,t'}) \geq 0,\ \forall\, i \in I, \forall\, t,t' \in T;$

(iii) $C'_{i,t} \geq 0,\ \forall\, i \in I, \forall\, t \in T,$

(iv) $\gamma_t > 0, \forall\, t \in T.$

Without sampling errors, we should look for marginal costs that satisfy properties above without γ_t. We treat γ_t as unknown and let the linear program look for the set of $\{\gamma_t, C'_{i,t}\}_{\forall\, i \in I, \forall\, t \in T}$ that rationalizes the data with the model. The idea is to ask if there are plausible sampling errors in the estimated aggregate $\hat{Q}_t$ and $p_t\hat{F}_t$ that would make the micro data consistent with the model. Furthermore, when more information (or modeler-defined judgement) of direction or range of the sampling errors is available, we can easily add bounds on the sampling errors to the constraints.[1]

In the linear program specified above, γ_t counts the ratio of sampling errors in total revenue and total input. It increases the bandwidth of the two variables and gives more flexibility to the constraints on marginal costs. Compared to the basic model, we would expect lower rejection rates of the model when sampling error is allowed. Meanwhile, estimates of the sampling errors $\{\gamma_t\}_{\forall\, i \in I, \forall\, t \in T}$ associated with the corresponding rejections to the model inform us about the sensitivity of the tests to sampling errors. In our application below, we compare results for the same sample with and without sampling errors.

3.2. Distance to the Model and Statistical Tests

Following the logic of sampling error in Section 3.1, relaxing the constraints results in lower rejections to the model. Building on the marginal-cost-consistency methods described in

[1] For example, if the modeler suspects $\beta > 1$, concavity of F implies $\alpha < \beta$, so $\gamma < 1$; the opposite would follow if $\beta < 1$.

Afriat (1972), Diewert (1973), and Varian (1985), we can gauge the distance of the revealed marginal costs in our tests to those that are consistent with the TOC model. Similar to Varian's approach of finding a minimal perturbation of the budget constraints that would make observed choices consistent with GARP, we can find a minimal adjustment to marginal costs needed to turn a rejection of the model to acceptance.

We implement this method by adding adjustment factors to marginal costs in the common ratio property, but not the co-monotone property. The idea is that the marginal costs in the co-monotone property describe the true convexity of the cost function, but firms may treat the marginal costs as being different in their objective function. The adjustment factors are constructed in a way to guarantee that data would always pass the model. We use a linear program to find the minimal magnitude of the adjustment, which is the minimized distance from the revealed marginal costs to those that would be consistent to the model. We denote them as revealed marginal costs and model-consistent marginal costs below, respectively. Based on these solutions, we then derive Kolmogorov-Smirnov and chi-squared tests to inform statistical acceptance/rejection of the model.

We use the following quadratic program:

$$\min_{C'_{i,t},\delta_{i,t}} \sum_t \sum_i \delta_{i,t}^2$$

Subject to:

(i) $\frac{p_t \mathrm{F}_t(Q_t) - Q_t(C'_{i,t} + \delta_{i,t})}{q_{i,t}} = \frac{p_t \mathrm{F}_t(Q_t) - Q_t(C'_{j,t} + \delta_{j,t})}{q_{j,t}} \geq 0, \ \forall\, i, j \in I, \forall\, t \in T;$

(ii) $(q_{i,t} - q_{i,t'})(C'_{i,t} - C'_{i,t'}) \geq 0 \ \ \forall\, i \in I, \forall\, t, t' \in T;$

(iii) $C'_{i,t} \geq 0 \ \ \forall\, i \in I, \forall\, t \in T.$

$\delta_{i,t}$ is the minimum adjustment factor on marginal cost $C'_{i,t}$. Note that the $\delta_{i,t}$ appear only in condition (i), not (ii). Again, the intuition here is that the cost functions are convex (ii), but firms may make errors in their optimization which show up in their first-order conditions (i). Alternatively, the analyst has made an error in the modelling of the objection function, which also shows up in condition (i).

Constraints (i), (ii) and (iii) guarantee that the set $\{\delta_{i,t}, C'_{i,t}\}$ satisfies the common-ratio property, co-monotone property, and nonnegativity constraint. By construction, such solutions always exist.[2] Hence, we can identify and quantify the minimal squared adjustment factors $\{\delta_{i,t}\}$, which are the minimal distances between the revealed marginal costs to the model-consistent marginal costs.

Statistical Tests

Taking the minimal distance found above, we can conduct a Kolmogorov-Smirnov (KS) test of the null hypothesis that the data are consistent with the model. Denote the set of marginal costs that are consistent with the model as $\{\widetilde{mc_{i,t}}\}_{i \in I, \forall t \in T}$ (model-consistent marginal costs). The model-consistent marginal costs can be obtained from the linear program in this section as $C'_{i,t} + \delta_{i,t}$. Denote the revealed marginal costs of an observed data set as $\{\widehat{mc_{i,t}}\}_{i \in I, \forall t \in T}$. The revealed marginal costs are obtained in the linear program as $C'_{i,t}$.

The two-sample KS test directly compares the distance between the cumulative probability function (CDF) of two sample variables and checks if the two samples are from the same distribution. The empirical distance function is specified as $D_{n,m} = sup_x |F_{1,n}(x) - F_{2,m}(x)|$, which represents the supremum of the distance between the CDF of sample 1 with n observations and the CDF of sample 2 with m observations. In our case, sample 1 consists of the model-consistent marginal costs, and sample 2 the revealed marginal costs. The sample size for both samples is $I * T$. $D_{n,m}$ is a vector consisting of the distance between the two CDFs at each value of the sample variable represented by x, which in our case is the marginal cost. We can take small intervals on the domain of marginal costs, obtain values of the two CDFs, and find the maximum distance of the two CDFs. The null hypothesis is rejected at level α if the maximum distance is larger than the critical value, that is $D_{n,m} > c(\alpha) * \sqrt{\frac{m+n}{m*n}}$, at critical level α.

Alternatively, we can assume the model-consistent marginal costs follow a log-normal distribution $N(\mu, \sigma^2)$ with the lower limit zero. Under the null hypothesis that an observed data set

---

[2] This is because the adjustment factors expand the domain of marginal costs to all real numbers. As there is no convexity constraints on the adjustment factors (i.e. no co-monotone constraint), adjustment factors can always be found to make the common-ratio properties be satisfied. Note that it would not do to incorporate the adjustment into all equations. That would simply be the same as the original model. If there are no numbers $C'_{i,t}$ satisfying (i)-(iii), then there are no numbers $(C'_{i,t} + \delta_{i,t})$ either.

is consistent with the model, revealed MCs would converge to the distribution of model-consistent MCs in the limit. Hence, $z_{i,t} = \frac{\log(\widehat{mc_{i,t}})-\log(\widetilde{mc_{i,t}})}{\sigma}$ follows a standard normal distribution. And we can easily obtain $z_{i,t}$ from the program, given that $\widehat{mc_{i,t}} = C'_{i,t}$ and $\widetilde{mc_{i,t}} = C'_{i,t} + \delta_{i,t}$. As a result, $S = \sum_{t=1}^{T}\sum_{i=1}^{I} {z_{i,t}}^2$ follows a chi-squared distribution with $T * I$ degrees of freedom. With a large sample, we can substitute the sample variance for the population variance. When S is larger than the critical value of a chi-squared distribution, we can reject the null that the data is consistent with the TOC model statistically.

3.3. Measurement Error in *q*

In Section 3.2, we considered distance to the model in the space of marginal costs as they show up in Condition (i), marginal cost consistency. An alternative is to consider distance to the model in the space of inputs $q_{i,t}$. If we allow those to be measured with error, then we can frame this approach as asking, how large would measurement error in inputs have to be for it to explain any rejections of the model?

In this case, we can again minimize $\sum_t \sum_i \delta_{i,t}^2$, but with $q_{i,t}$ replaced by $(q_{i,t} + \delta_{i,t})$ and similarly Q_t replaced by $(Q_t + \sum_i \delta_{i,t})$ everywhere in the model. If we denote the model-consistent inputs as $\tilde{q}_{i,t} = \left(q_{i,t} + \delta_{i,t}\right)$ and similarly the sum $\tilde{Q}_t = \left(Q_t + \sum_i \delta_{i,t}\right)$, we can write this more succinctly as finding the model-consistent inputs $\tilde{q}_{i,t}$ that are closest to the observed inputs. This involves the non-linear program:

$$\min_{C'_{i,t},\tilde{q}_{i,t}} \sum_t \sum_i \left(\tilde{q}_{i,t} - q_{i,t}\right)_{i,t}^2$$

Subject to:

(i) $\frac{p_t \mathrm{F}_t - \tilde{Q}_t C'_{i,t}}{\tilde{q}_{i,t}} = \frac{p_t \mathrm{F}_t - \tilde{Q}_t C'_{j,t}}{\tilde{q}_{j,t}} \geq 0, \ \forall\, i,j \in I, \forall\, t \in T;$

(ii) $\left(\tilde{q}_{i,t} - \tilde{q}_{i,t'}\right)\left(C'_{i,t} - C'_{i,t'}\right) \geq 0 \ \ \forall\, i \in I, \forall\, t,t' \in T;$

(iii) $C'_{i,t} \ \geq 0 \ \ \forall\, i \in I, \forall\, t \in T.$

(iv) $\tilde{q}_{i,t} \ \geq 0 \ \ \forall\, i \in I, \forall\, t \in T.$

Note the non-linear constraints in Expressions (i) and (ii). The basic idea here is to find *some* set of inputs that are consistent with the outputs and the model restrictions, but to find those inputs closest to the observed data. This approach has the advantage of a clear structural interpretation in terms of measurement error and of consistently incorporating the error into all relevant points in the model.

3.4. Dynamic Resources and Other Games

Our basic model in Section 2 pertains to the static average-return game with Nash behavior. Banzhaf and Liu (2016) further show these results can be extended to the case of conjectural variations (rather than Nash behavior) suggested by Cornes and Sandler (1983). They also show they can be modified to apply to the average cost (rather than average returns) game, where agents choose outputs and pay the average costs. Such problems are relevant to many problems involving the division of joint costs, such as telephony.

In Appendix B, we further show that our results apply to dynamic resources, where the tragedy of the commons applies to the dissipation of the in-situ value of leaving resources in place (Clark 1980, Levhari and Mirman 1980). In general, the dynamic model requires additional restrictions. However, as we show in the appendix, our basic model of Section 2 applies whenever firms treat the *in situ* value of the stock as zero or, alternatively, as proportionate to their catch shares.

4. Empirical Application

We apply our test to the Norwegian whitefish fishery using data for the period 1998 to 2007. The setting is fitting for two reasons. First, open-access fisheries are a classic example of the tragedy of the commons. Second, this particular fishery experienced a management change that strengthened property rights and thereby reduced tragedy of the commons incentives over the period studied, such that we would expect the tragedy of the commons model to fit the data better in the first part than the second part, allowing for a comparative test of two regimes.

In the remainder of this section, we further describe the Norwegian fishery and the data available.

4.1. The Norwegian Ground Fishery

Norway has the largest fishing industry in Europe. Its most valuable fishery is whitefish, with cod, haddock and saithe (Atlantic Pollock) being the most important species. Norway's white-fish fishery is biologically separate from other major fisheries, so output from the fishery $F(Q)$ can be modeled in isolation as a single resource. The fleet targeting whitefish comprises various vessel groups of different sizes and gear. Trawlers are relatively large vessels, with lengths ranging from 28 to 76 meters, and fish in deeper waters. The coastal fleet comprises smaller vessels using a variety of gear such as long lines, troll nets and Danish seine. Our sample contains only the coastal fleet. The management system requires that each fishing vessel is separately owned by an operator, so vessels can be taken as firms in our model.

In 1989 a total allowable catch (TAC) quota was set for the whole whitefish fishery, with the TAC divided between the trawler fleet and the coastal fleet. In 1990, a non-tradable individual vessel quota (IVQ) system was theoretical introduced to the Norwegian coastal fleet. To ensure that the allocated quotas were fished within the coastal vessel group, an "overbooking system" was introduced in 1991 where the sum of the individual vessels' quotas were higher than the TAC for the vessel group. As the overbooking was substantial, the IVQ system essentially was not binding, making the management more like a regulated restricted access system (RRA) than a true IVQ system. From the perspective of our theoretical model, we view this period as preserving the open access regime, with some restrictions on technological inputs and total catch, but with no individual limits on catch (or effort) and with incentives promoting a race to fish. Our data (described below) begin in 1998, during this regime.

In 2003, the quota for the coastal fleet was divided into four groups by vessel length. Groups no longer needed to compete across size categories. This appears to have helped the small vessels as a group. However, the sum of the individual quotas still exceeded the TAC (group quota), so though firms theoretically could catch all their quota, they still had to compete with other vessels of the same size class to reach the limit. Moreover, there was no guarantee they would get any quota. Effectively, the individual quotas were upper-bound constraints.

Finally, in 2004, overbooking ended for vessels above 15 meters. Additionally, these large vessels now could combine quotas from several vessels onto one, thereby introducing transferability into the system. Thus, the regime for larger coastal vessels transformed to a truly binding IVQ system in 2004, while it remained an RRA system for smaller vessels. Hannesson (2013), Standal et al. (2016) and Cojocaru et al. (2019) provides further information about the fishery and the development of the management system.

In sum, from 1998 to 2002, all vessels in our data set were under an RRA regime. After 2003, larger coastal vessels transitioned into an IVQ regime while the small vessels were still under an RRA regime. In between, 2003 was something of a transition year. Small vessels and large vessels were given separate group quotas, but still competed within group, a problem that may have been especially severe for small vessels.

This change in property rights regimes affords an opportunity to apply our test of the tragedy of the commons using a difference-in-differences design. We expect higher rejection rates for big coastal vessels for the 2003-7 period, relative to the 1998-2002 period, and relative to the corresponding difference for small vessels. In sensitivity analyses, we also consider omitting 2003.

4.2. Description of Data

The data for the Norway coastal fleet cover the period 1998 to 2007 and come from an annual random survey of vessels with only a sample of the registered active vessels being surveyed each year. Table 1 summarizes the data. The first row shows the sample size. The second row shows the total number of vessels registered in each year (population). The total sample comprises 1127 individual vessels from 1998 to 2007. Each vessel is identified with a unique ID. We have information on the length and weight of each vessel as well as on effort and other inputs, including days at sea, operating days (days at sea plus days working at port), fuel expenditure, labor compensation, and the average number of crew members operating the vessel.

With respect to outputs, we have vessel-year data on the total quantity landed and revenues received by species (cod, haddock, saithe and other whitefish), in tons and Norwegian Krone (NOK), respectively. However, our test only requires knowing the aggregate revenue. Thus, we first create an index by summing over fish species, then sum over vessels to obtain the total sample revenue for each year, $p_t\hat{\mathrm{F}}_t$. Then, we multiply the average sample revenue by number of total vessels in the population to obtain the aggregate revenue. Row 3 of Table 1 shows the total sample revenue. Row 4 converts this sample to an estimate of total revenue, multiplying the sample mean by the number of vessels operating. This is the value of output $p_t\hat{\mathrm{F}}_t$ used in our test. It shows some ups and downs followed by an upward trend after 2003. The next row similarly shows the trend in sampled catch in tons. The remainder of Table 1 offers additional details on the distribution of catch across vessels, by species and year. We offer these data for completeness, but only use Row 4 from this table in our empirical work.

Although it requires only annual aggregate revenue on the output side, our test requires micro-level data on the input side. Vessels are not necessarily sampled in each year and do not necessarily fish in all years anyway, so we have an unbalanced panel of vessel-level inputs. Also, reported zeros for an input indicates that these fields were left blank in the survey. Accordingly, we exclude vessels that reported both zero operating days and zero days at sea but positive labor, fuel or other operating expenses in the analysis. Table 2A shows raw data on inputs, including operating days, days at sea, person-years, labor compensation, and fuel expenditure.

4.3. Quantifying Effort

In taking the theoretical model to the data, a central modelling question is how to measure effort (or input) $q_{i,t}$ as a scalar, as required by the theoretical model. As measures of effort, we consider the following four proxies: operating days, imputed days at sea, imputed days at sea times vessel length (Length* Days), and an estimated scalar-valued function of effort based on multiple inputs. Of these, operating days, which includes days at sea as well as days processing and offloading in port, is the most straightforward proxy. Table 2B shows summary statistics for operating days as used in the model.

Our second measure is days at sea. Averaging over time, days at sea contains 81.3 fewer days fleet-wide than operating days, and there are 748 observations with positive operating days but zero reported days at sea. Since it is impossible to have zero days at sea when operating days and catch are positive, we treat these zeros as missing and replace them with imputed values when the associated operating days are positive. To impute these values, we use the following regression model:

$$(6) \qquad days\ at\ sea = \beta_0 + \beta_1 * operation\ days + \beta_2 * fuel\ expenditure.$$

We run the model in Equation (6) conditional on $operation\ days > 0$ and $days\ at\ sea > 0$, and use the predicted coefficients to estimate missing values of days at sea for observations with positive operating days. Table 3 gives the estimated regression coefficients from Equation (6) (Model 4), as well as alternatives. Model 1 estimates days at sea only as a fixed proportion of operating days; Model 2 adds fuel expenditure but continues to omit the constant. Models 3 and 4 are similar to 1 and 2 respectively, but include the constant term. Out-of-sample prediction comparisons (using leave-one-out validation) suggest that Model 4 has the best fit, with the exception of Model 5, which includes fixed effects. However, vessel fixed effects cannot be estimated for those vessels with insufficient data, making this an impractical choice. Thus, we choose Model 4 as it reflects a balance between accuracy and reducing missing observations. Based on this model, Table 2B shows annual data on imputed days at sea.

Our third measure of input uses these imputed days at sea times vessel length. Rescaling fishing time by measures of vessel size is a common practice when estimating fisheries production functions, as a better measure of overall inputs (Squires 1987; Huang and Smith 2014). Table 2B also reports annual values of this product.

Our fourth and final measure of input aggregates multiple input variables into a scalar-valued function. This too is a common practice in the fisheries literature (see McCluskey and Lewison 2008 for review and discussion). We adopt a straightforward method that serves our purpose. Suppose the production function for vessel i in year t is

$$\ln\left(Catch_{i,t}\right) = a + b * lnE_{i,t} + \lambda_t + e_{i,t}, \tag{7}$$

where λ_t is a dummy which captures year effects, such as different stock levels, and $E_{i,t}$ denotes the overall effort level for vessel i at year t, and is a sub-function of other inputs. In particular, let

$$\begin{aligned} \ln(E_{i,t}) = {} & \alpha_2 \ln(person\text{-}years_{i,t}) + \alpha_3 \ln(fuel\ expenditure_{i,t}) + \\ & \alpha_4 \ln(labor\ compensation_{i,t}) + vesselid_i, \end{aligned} \tag{8}$$

in which man-years denotes the labor input (measured at the day level) and labor compensation is the total payment to workers on the vessel and $vesselid_i$ is vessel level fixed effect that captures vessel length, tonnage, etc.

Substituting Equation (8) into (7), we estimate the combined model. Note, however, that we cannot separately identify b in Equation (7) from the alphas in Equation (8). Thus, we do not identify effort to scale. This is not problematic, however, because our test treats the cost of effort as a latent function, so any arbitrary change of scale in effort can be reconciled by an offsetting change in the scale of the cost function. The results of estimating this model are shown in Table 4.

Column 1 introduces the individual inputs in levels, whereas Column 2 does so in logs (as shown in Equation (8)). We use Column 2 in our analysis, as it has a better fit. Table 2B shows summary statistics for this estimated value.

4.4. Sampling Subsets of Data

Because, in our approach, rejections are all or nothing, the presence of only one firm behaving out of step with the other firms could result in rejecting the entire data set. Likewise, if cost functions shift over time, assuming they are constant could lead to false rejections. To sidestep these issues, we follow Carvajal et al. (2013) and repeatedly sample smaller subsets of data. Sampling the data allows us to consider rejection *rates* (percentage of data sets that do not conform to the tragedy of the commons model), rather than one single all-or-nothing conclusion. We follow Carvajal et al. (2013) and repeatedly sample smaller subsets of data. We divide the entire data set into multiple subsets, with each set consisting of N vessels and T consecutive years, where $N \in \{5, 10, 50, 100, 150\}$ and $T \in \{3, 6, 8, 10\}$. Then we separately test for consistency with the tragedy–of–the–commons model using each set. We randomly sampled 100 subsets from each N-by-T combination, giving us a reasonable estimate of the rejection rates for each combination. (To facilitate comparisons, we used the same subsample of data for each cell across models.)

4.5. Weighted Sampling and Property Rights Regime Comparison

As discussed in section 4.1, the evolution of property rights in the Norwegian fishery motivates splitting the data into the periods of the RRA regime (1998–2002) and the period of IVQs for the coastal vessels at least 15 meters in length (2003–2007). Accordingly, we cut the data into four cells using a 2x2 design; large coastal vessels (≥15 meters long) and small (<15 m), before the IVQ regime (1998–2002) and after (2003–2007).

It is worth noting that, though we sub-sample by vessel size in this exercise, in the common-ratio properties for each group of each year, we keep the total input Q_t and output $F_t(Q_t)$ across all vessels. That is, behavior by all vessels (regardless of length) still affects the optimal behavior of any one vessel.

In this unbalanced panel for the Norwegian coastal fleet, due to the administration of a random survey, there are fewer observations of surveyed vessels in earlier time periods (before

2003) than later (after 2003). When we sample subsets as described in Section 4.4 with no restrictions (where each vessel has an equal probability to be selected), the sets sampled in later periods will contain more data points than those from earlier periods. Given the nature of our test, more data points create more constraints, which automatically yields higher rejections holding all other things equal. Hence, to make sure the gap in rejection rates per group is attributed to behavioral difference under different management regimes, rather than the difference in the number of observations in the samples, we employ weighted sampling to generate comparable samples for each group.

Weighted sampling is implemented by redistributing sampling probabilities among vessels in later periods (2003-2007). Sampling probabilities for vessels with more observations (3 and 4 data points in periods 2003-2007) are reduced, and the reduced probabilities are added to vessels with fewer observations (1 and 2 data points), with the total probability always summing to one. The largest adjustment of the probability of a vessel is less than 0.0002, while the original probability of a vessel being sampled is around 0.00116, so the adjustment is less than 17%. After weighted sampling, the maximum difference in the number of observations between the groups (before vs. after) is less than 0.2% (difference in observations divided by total observations in subsample sets). In our 2x2 design, our weighted sampling ensures that the big-after and big-before groups have similar numbers of observations, as do the small-after and small-before groups. This helps to balance the number of observations among groups to generate credible difference-in-difference results.

As discussed in Section 4.1, data generated from the IVQ regime is not expected to be consistent with the tragedy of the commons model, especially for large vessels. Accordingly, we first take the difference of rejections between the big-after and big-before groups and likewise for the small-after and small-before groups. Finally, we take the difference-in-differences, to infer the effects of the change in property rights regime. We expect the after-before difference for big vessels will be higher than those for small vessels.

5. Results

In this section, we present the results of our tests. We first present results of the basic tests as described in Section 2. We then present results with sampling errors (Section 3.1) and statistical tests based on distance from revealed marginal costs to model-consistent marginal costs (Section 3.2). Finally, we present tests using our difference-in-differences design.

5.1. Results of Test Pooling all Data

Tables 5A-5D present results using the basic test of Section 2, using four respective proxies of effort: operating days, imputed days at sea, imputed days at sea times length, and estimated total effort. Each cell in the tables shows the rejection rate for a sample of 100 data sets for N vessels and T consecutive years, for varying N and T. For small N and T, we generally cannot reject the tragedy of the commons model in most samples. Note, however, that the rejection rates generally are increasing in N (moving down the rows) and T (moving to the right across columns). Indeed, when more than 100 vessels are considered for longer than 6 years, the rejection rates approach one. This trend is necessary, mechanically, as the number of equations and inequalities to satisfy is increasing in these parameters, so exceptions to this rule are due to random sampling. More substantively, the trend also is consistent with the idea that, as we increase T, we risk pooling different cost functions as well as data from the period after the property rights reform, when the TOC model is unlikely to apply. Overall, these results indicate that the behavior of vessels/fishermen in our sample cannot be explained by the TOC model when a large number of observations are included.

Additionally, we test consistency with the model with sampling errors (as discussed in Section 3.1). The boundaries on sampling errors we adopted is [-5%, 5%]. That is, we restrict the multiplier γ_t to be between [0.95, 1.05]. We are only able to apply narrow boundaries to our sample data from Norwegian ground fishery due to the large number of missing values in the sampled data.[3] Notice that the adjustment factor functions as a multiplier on total revenue. Given that the average revenue in our sample is 1.4 million NOK (around 166,000 USD) per year per vessel, this bandwidth allows for an average adjustment to the revenue of 67,000 NOK (around 8,000 USD) per year per vessel. That amount is more than the average cost of fuel expenditure per year per vessel, so it is not negligible.

[3] Our unbalanced panel data of Norwegian ground fishery has 79.3% of data points missing. The amount of missing substantially reduces nonempty constraints in our test, which makes it easy to find marginal costs that are consistent with the model. Allowing for a larger adjustment to the total revenue makes the tests even less stringent and reduces the rejection rates towards zero. For instance, all rejection rates are zero when the boundary is 10% in our case.

Tables 6A-6D present results using the test with sampling errors (Section 3.1). As we would expect with added flexibility, rejections to the TOC model allowing for sampling errors are slightly lower than those in the basic model (comparing like cells). But the previous patterns remain. First, rejection rates still increase in N and T. Second, when more than 100 vessels are considered for longer than 6 years, the rejection rates still approach one. This result provides additional support for the conclusion that behavior of vessels/fishermen in our sample cannot be explained by the TOC model when a fair number of observations are included.

We also conducted the KS test of Section 3.2 to the entire data set. For all four measures of effort, we reject the tragedy of the commons model with the pooled data with p-values < 0.01. Results from this tests confirm our observation from the rejection rates in Tables 5 and 6.

5.2. Results Comparing Property Rights Regimes

Recall that all vessels operated under RRA before 2003. Throughout the period (1998-2007) in our sample, a TAC for all participants was in place, but in 2003 the quota was distributed to groups based on vessel length. After 2003, small vessels remained operating under a total allowable catch and the RRA regime, while big vessels transitioned to an IVQ regime. This make the small vessels a good control group for the big vessels: whereas there is competition among vessels under a group quota, competition among big vessels is reduced under the property-rights based management of IVQs. The effectiveness of the property-rights approach of IVQs over the non-property-rights based approach of RRA drives the difference-in-differences results in our empirical study.

Table 7A – 7D present rejection rates per group using the weighted sampling described above in Section 4.5. The results indicate that, after the reform, big vessels incur a higher increase in rejection rates of the TOC model than small vessels. That implies the IVQ regime generates more fishing behavior inconsistent with the tragedy of the commons model. In other words, the IVQ regime nudges fishing behavior away from Nash more effectively than does RRA, as one would expect.

Note that after we split the data into four groups, there are fewer observations to sample from per group. Because the weighted sampling only controls for the *difference* in the number of observations of each paired group (before vs. after), but not the *magnitude* of observations in samples, the levels of rejection rates are sensitive to the number of observation in the respective subgroups, but the difference and difference-in-difference results do reflect the overall change in management regimes and are more stable.

We also replicated these tests omitting 2003, which was a transition year and arguable was different from the subsequent 2004-7 period, when large vessels were under the TAC. Our results are qualitatively similar using this approach. They are available upon request.

Interestingly, looking only at small vessels, we observe a decrease in rejection rates in the 2003-7 period. Taken in isolation, this suggests that the behavior of small vessels actually moved closer to the Nash Tragedy of the Commons behavior after 2003. One possible explanation for this finding is an induced race to fish among small vessels after securing a shared right for small vessels as a group but without assigning individual rights. Table 8 compares the number of small and big vessel across years. It shows that there is a marked increase in the total number of small vessels starting in 2003, whereas there is not much change in the number of big vessels. Even with a slight decline in average fishing effort in all vessels after 2003, the increase in the number of small vessels still leads to an increase in the total effort of the small-vessel group. The increased number of participants and increased total effort move the collective behavior of small vessels closer to Nash. New entry in small vessels may have been induced by increased economic rent after the division of the quota. Perhaps before 2003, under the TAC for all vessels, small vessels could not compete with big vessels in the race to fish.[4] After 2003, separating out the TAC for the small-vessels reduced the competition from big vessels and secured a potential economic rent. However, without individually assigned property rights to quotas, that potential rent attracted new entrants and spurred the race to fish. This interpretation is in line with the finding in Homans and Wilen (1997) that certain types of non-property-rights-based management may induce a race to fish. It also is consistent with the findings in Kroetz et al. (2015) that policy with good social objectives can reduce overall economic efficiency and rents in fisheries.

Table 9 shows the results of allowing minimal behavioral errors (Section 3.2). For each of the four measures of effort, it shows the mean squared error $\delta_{i,t}^2$ per cell, an adjusted mean squared

[4] Technically, our model captures the incentives even for small vessels with little market power in manipulating resource rents. However, in practice, it may be that with small costs of optimizing it did not make sense for small vessels to fully consider the incentives under Nash competition until the quota was divided.

error per thousand *constraints* to be satisfied, which we prefer,[5] and the p-value for the KS test. To gauge the scale of these estimated errors, the mean marginal cost is about 4.5 when effort is measured by operating days, so these errors are fairly small. This scaling differs by measure of effort making comparisons difficult, but, across measures, the mean absolute value of the errors is about 5% of marginal costs, the mode is 0%, and the 90th percentile error is an error of 10-17%. Comparing across vessel sizes and property rights regimes for any one measure of effort, we see a notable increase in the errors and, to some extent, the probability of rejecting the model in the "after" period relative to the "before" period, as we would expect. The difference in these differences across vessel sizes is not as clear as with the rejection rates. However, as a rule p-values cannot be meaningfully differenced across models. Focusing on our preferred measure of the adjusted MSE, we see greater increases for the large vessels, as we would expect.

6. Conclusion

Work to date on testing the tragedy of the commons has focused either on policy outcomes involving the state of shared resources or, when using behavioral data, has relied on highly structural models involving numerous maintained assumptions. Drawing on applications of revealed preference theory to behavioral data, such as work by Carvajal et al. (2013) on the Cournot model, we derive non-parametric tests of the tragedy of the commons using minimal behavioral assumptions. Additionally, we present methods to account for the sampling errors in aggregate output and input data, and to gauge the distances to the model as well statistical tests based on the distances.

We apply this new test to the Norwegian groundfish fishery. Overall, we find the behavior of individual fisherman/vessel of the Norwegian Coastal Fishery does not conform to the model of the tragedy of the commons. More importantly, we also find that rejection rates are larger after property rights reforms, especially for the large vessels that received stronger property rights. Moreover, using a distance-based metric, we find that behavior moves further from behavior associated with the pure tragedy of the commons model after the property rights reforms. Our results suggest that Norwegian policy has changed behavior and, presumably, ameliorated the commons

[5] For example, if there are N vessels and T years of data, and if there were no missing data, there would be NT cells used as the denominator for the simple mean squared error, but $N(T^2-T)/2 + T(N^2-N)/2 = NT(N+T-2)/2$ constraints used as the denominator for the adjusted mean squared error. Our actual calculation accounts for missing values in the formula.

problem at least for large coastal vessels.

Our model and approach allow for comparative work on the behavioral consequences of policy interventions to govern common-pool resources. Economists often imagine a stylized first-best policy to ration access to the commons as an optimal total quota that is divided among individual participants, with perfect security and transferability of the property right. However, real-world policies are configured in a myriad ways that differ from theoretical first-best policies to address the commons problem. Do some policy configurations move behavior away from the tragedy of the commons more than others? In fisheries, rights-based systems differ along dimensions of the security of the property right, the length of term, transferability, and a number of other restrictions that often come about as political compromises to address community or industry concerns (Asche et al 2018). The same species of fish that we analyze in this paper are regulated with very different rights-based systems in Iceland, Canada, and the United States that differ along these dimensions. Our model and distance-based metric have the potential to examine whether these different rights-based fisheries policies induce more or less commons-like behavior.

Our approach can also be applied to other common-pool resources whenever firm-level data on inputs are available. Candidates include clearcutting under different governance structures; grazing livestock on commons land; pumping groundwater; oil, gas, and other mineral extraction; and telephony and other cost-sharing problems.

References

Afriat, S.N. 1967. "The Construction of a Utility Function from Demand Data." *International Economic Review* 8: 67-77.

Afriat, S.N. 1972. "Efficiency Estimation of Production Functions." *International Economic Review* 13(3): 568-98.

Asche, Frank, Taryn M. Garlock, James L. Anderson, Simon R. Bush, Martin D. Smith, Christopher M. Anderson, Jingjie Chu, Karen A. Garrett, Audun Lem, Kai Lorenzen, Atle Oglend, Sigbjørn Tveteras, and Stefania Vannuccini. 2018. "Three Pillars of Sustainability in Fisheries." *Proceedings of the National Academy of Sciences* 115(44): 11221-25.

Ayres, Andrew B., Kyle C. Meng, and Andrew Plantinga. 2019. "Can Property Rights Alleviate the Problem of the Commons? Evidence from California Groundwater Permits." Working Paper.

Balthrop, Andrew T., and Kurt E. Schnier. 2016. "A Regression Discontinuity Approach to Measuring the Effectiveness of Oil and natural Gas Regulation to Address the Common-Pool Externality." *Resource and Energy Economics* 44: 118-38.

Banzhaf, H. Spencer, and Yaqin Liu. 2016. "Notes to Nonparametric Test of the Tragedy of the Commons." Working Paper.

Birkenbach, Anna M., David J. Kaczan, and Martin D. Smith. 2017. "Catch Shares Slow the Race to Fish." *Nature* 544: 223-26.

Brozović, Nicholas, David L. Sunding, and David Zilberman. 2010. "On the Spatial Nature of the Groundwater Pumping Externality." *Resource and Energy Economics* 32(2): 154-64.

Burlig, Fiona, Louis Preonas, and Matt Woerman. 2019. Spatial Externalities in Groundwater Extraction: Evidence from California Agriculture. Working Paper.

Carvajal, Andres, Rahul Deb, James Fenske, and John K. H. Quah. 2013. "Revealed Preference Tests of the Cournot Model." *Econometrica* 81: 2351-79.

Chambers, Christopher P., and Federico Echenique. 2016. *Revealed Preference Theory*. Cambridge University Press.

Ciriacy-Wantrup, S.V., and Richard C. Bishop. 1975. "'Common Property' as a Concept in Natural Resource Policy." *Natural Resources Journal* 15: 713-27.

Clark, Colin W. 1980. "Restricted Access to Common-Property Fishery Resources: A Game-Theoretic Analysis. In *Dynamic Optimization and Mathematical Economics*, ed by Pan-Tai Liu, pp. 117-32. Springer.

Cojocaru, Andreea, Frank Asche, Ruth B. Pincinato and Hans-Martin Straume. 2019. "Where Are the Fish Landed? An Analysis of Landing Plants in Norway." *Land Economics* 95(2): 246-57.

Cornes, Richard, and Todd Sandler. 1983. "On Commons and Tragedies." *The American Economic Review* 73(4): 787-92.

Cornes, Richard, and Todd Sandler. 1996. *The Theory of Externalities, Public Goods, and Club Goods*. Cambridge University Press, 2nd ed.

Costello, Christopher, Steven D. Gaines, and John Lynham. 2008. "Can Catch Shares Prevent Fisheries Collapse?" *Science* 321: 1678-1681.

Dasgupta, Partha S., and Geoffrey M. Heal. 1979. *Economic Theory and Exhaustible Resources*. Cambridge University Press.

Diewert, W. Erwin. 1973. "Afriat and Revealed Preference Theory." *Review of Economic Studies* 40: 419-25.

Felthoven, Ronald G., William C. Horrace, and Kurt E. Schnier. 2009. "Estimating Heterogeneous Capacity and Capacity Utilization in a Multi-Species Fishery." *Journal of Productivity Analysis* 32: 173-89.

Gordon, H. Scott. 1954. "The Economic Theory of a Common-Property Resource: The Fishery. *Journal of Political Economy* 62:124-42.

Hands, D. Wade. 2014. "Paul Samuelson and Revealed Preference Theory." *History of Political Economy* 46(1): 85-116.

Hannesson, Rögnvaldur. 2013. "Norway's Experience with ITQs." *Marine Policy* 37: 264-69.

Hardin, Garrett. 1968. "The Tragedy of the Commons." *Science* 162: 1243-48.

Huffaker, Ray G., and James E. Wilen. 1991. "Animal Stocking under Conditions of Declining Forage Nutrients." *American Journal of Agricultural Economics* 73(4): 1213-23.

Huang, Ling, and Martin D. Smith. 2014. "The Dynamic Efficiency Costs of Common-Pool Resource Exploitation." *American Economic Review* 104: 4071-4103.

Kirkley, James, Catherine J. Morrison Paul, and Dale Squires. 2002. "Capacity and Capacity Utilization in Common-Pool Resource Industries." *Environmental and Resource Economics* 22: 71-97.

Koch, Caleb M., and Heinrich H. Nax. 2017. "Rethinking Free Riding and the Tragedy of the Commons." Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3075935

Kroetz, Kailin, James N. Sanchirico, and Daniel K. Lew. 2015. "Efficiency costs of social objectives in tradable permit programs." *Journal of the Association of Environmental and Resource Economists*2(3): 339-66.

Levhari, David, and Leonard J. Mirman. 1980. "The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution." *Bell Journal of Economics* 11(1): 322-34.

Libecap, Gary D., and Steven N. Wiggins. 1984. "Contractual Responses to the Common Pool: Prorationing of Crude Oil Production." *American Economic Review* 74(1): 87-98.

McCluskey, Shannon M., and Rebecca L. Lewison. 2008. "Quantifying Fishing Effort: A Synthesis of Current Methods and their Applications." *Fish and Fisheries* 9: 188-200.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.

Ostrom, Elinor. 2009. "Building Trust to Solve Commons Dilemmas: Taking Small Steps to Test an Evolving Theory of Collective Action." In *Games, Groups, and the Global Good*, ed. by Simon Levin, pp. 207-28. Springer.

Sen, Amartya K. 1966. "Labour Allocation in a Cooperative Enterprise." *Review of Economic Studies* 33: 361-71.

Squires, Dale. 1987. "Fishing effort: Its testing, specification, and internal structure in fisheries economics and management." *Journal of Environmental Economics and Management* 14(3): 268-282.

Standal, Dag, Signe Annie Sønvisen, and Frank Asche. 2016. "Fishing in Deep Waters: The Development of a Deep-Sea Fishing Coastal Fleet in Norway." *Marine Policy* 3: 1-7.

Stavins, Robert N. 2011. "The Problem of the Commons: Still Unsettled after 100 Years." *American Economic Review* 101(1): 81-108.

Varian, Hal R. 1982. "The Nonparametric Approach to Demand Analysis." *Econometrica* 50(4): 945-73.

Varian, Hal R. 1984. The Nonparametric Approach to Production Analysis. *Econometrica* 52(3): 579-97.

Varian, Hal R. 1985. "Non-Parametric Analysis of Optimizing Behavior with Measurement Error." *Journal of Econometrics* 30: 445-58.

Weitzman, Martin. 1974. "Free Access vs. Private Ownership as Alternative Systems for Managing Common Property." *Journal of Economic Theory* 8: 225-34.

Table 1. Summary Statistics for Selected Output Variables

Variable		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Obs.		307	321	328	323	316	279	321	306	317	359
Population		1193	1143	1081	1063	1230	1441	1342	1131	1165	1290
Sampled annual value (100 mil. NOK)		3.61	3.67	3.67	3.91	3.98	4.54	4.61	4.68	6.58	7.40
Total annual value (100 mil. NOK)		17.64	14.91	13.83	15.60	14.33	12.58	13.55	14.65	19.62	19.30
Sampled annual harvest (10 million kg)		4.17	4.62	4.94	5.31	5.81	6.64	7.84	8.23	8.43	9.25
Cod	Mean	77.7	55.2	45.0	48.3	52.2	51.5	59.4	72.0	85.4	73.7
(thousand kg)	SD	87.2	60.3	53.6	51.2	38.5	38.3	45.4	63.2	72.3	66.6
	Min	0.1	0.9	0.6	0.2	0.1	0.2	0.1	0.2	0.3	0.0
	Max	471.4	411.1	581.8	334.6	332.6	299.3	294.6	452.0	444.4	451.3
Haddock	Mean	19.8	10.7	9.0	11.4	12.7	12.6	11.4	16.7	17.7	21.4
(thousand kg)	SD	38.3	21.9	19.7	14.3	26.9	32.7	21.3	30.4	28.2	38.7
	Min	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Max	204.3	188.1	211.3	92.4	251.3	416.2	158.5	260.5	185.0	310.8
Saithe	Mean	29.9	26.3	22.8	24.7	19.7	23.2	22.8	31.9	50.1	47.3
(thousand kg)	SD	68.9	49.5	32.9	42.6	37.8	33.3	38.0	68.5	101.6	101.6
	Min	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Max	574.1	418.7	251.7	420.0	321.1	197.3	199.2	716.4	873.8	943.7
Other	Mean	70.4	58.6	91.3	51.9	40.5	41.1	32.8	45.3	61.9	71.7
(thousand kg)	SD	248.2	212.3	302.6	178.1	131.9	94.7	77.7	110.4	162.4	263.9
	Min	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Max	1,807.2	1,859.2	2,203.4	1,864.4	1,409.4	644.3	673.4	899.4	2,014.3	2,482.1

Table 2A. Summary Statistics for Selected Input Variables (Raw Data)

Variable		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Obs.		69	72	80	76	71	279	321	306	317	359
Operating days	Mean	268.2	262.0	268.5	253.8	244.2	213.3	193.8	220.9	227.4	210.1
	SD	32.6	41.1	41.1	45.2	44.0	54.7	51.7	56.2	57.0	53.6
	Min	204.0	176.0	190	107	146	99.0	83.0	90.0	93.0	90.0
	Max	338.0	364.0	348	338	342.0	354.0	342.0	345.0	355.0	338.0
Days at sea	Mean	219.4	211.4	198.3	175.5	178.2	168.7	168.8	178.3	189.5	168.9
	SD	33.2	40.0	50.1	42.8	46.6	46.9	46.0	58.7	56.2	53.9
	Min	152.0	117.0	60.0	50.2	95.0	72.0	77.0	55.0	72.0	68.2
	Max	295.0	322.0	343.0	335.0	287.0	336.0	324.0	330.0	345.0	325.0
Person years	Mean	2.3	2.2	2.1	2.2	2.1	2.2	2.1	2.3	2.4	2.4
	SD	1.8	1.8	1.8	1.6	1.6	1.4	1.3	1.5	1.5	1.5
	Min	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Max	12.0	12.0	12.7	11.0	12.6	10.7	8.1	10.0	8.1	9.0
Labor compensation (thousand NOK)	Mean	637.3	607.6	574.8	652.3	593.7	511.2	607.4	772.3	1025.8	1015.9
	SD	799.9	808.9	791.9	821.6	592.5	480.4	562.8	721.6	937.6	979.2
	Min	65.5	81.5	65.8	63.1	109.3	104.1	108.0	149.1	141.5	158.2
	Max	5,161.4	6,658.9	5,930.7	6,151.7	4,918.5	3,906.7	4,606.4	4973.9	6920.2	7184.6
Fuel expenditure (thousand NOK)	Mean	47.9	52.3	80.6	70.6	59.8	59.7	72.6	108.0	135.5	121.6
	SD	73.0	91.9	161.3	127.3	108.1	92.6	97.9	163.7	177.8	194.1
	Min	3.0	3.4	1.5	4.6	3.2	1.3	3.1	6.9	10.2	9.6
	Max	539.5	745.7	1,405.7	1,458.6	1,066.7	1,113.5	937.7	1610.0	1605.5	1623.6

Table 2B. Summary Statistics for Selected Input Variables (As used in Analysis)

Variable		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Obs.		69	72	80	76	71	279	321	306	317	359
Operating days	Mean	258.2	262.0	268.5	253.8	244.2	213.3	193.8	220.9	227.4	210.1
	SD	32.6	41.1	41.1	45.2	44.0	54.7	51.7	56.2	57.0	53.6
	Min	204.0	176.0	190.0	107.0	146.0	99.0	83.0	90.0	93.0	90.0
	Max	338.0	364.0	348.0	338.0	342.0	354.0	342.0	345.0	355.0	338.0
Imputed days at sea	Mean	217.4	211.4	198.3	175.5	178.2	168.7	168.8	178.3	189.5	169.0
	SD	33.2	40.0	50.1	42.8	46.6	46.9	46.0	58.7	56.2	53.9
	Min	152.0	117.0	60.0	50.2	95.0	72.0	77.1	55.0	72.0	68.2
	Max	295.0	322.0	343.0	335.0	287.0	336.0	324.6	330.0	345.0	325.0
Length times Imputed days at sea	Mean	4169.2	4067.3	3748.1	3248.1	3197.8	2200.5	2237.6	2434.3	2605.8	2377.7
	SD	1261.4	1449.2	1713.7	1312.7	1377.6	1090.4	1146.4	1349.6	1260.3	1247.1
	Min	2133.6	1772.6	877.8	707.8	1459.2	696.0	672.0	581.9	816.4	606.6
	Max	7707.8	8826.0	9415.4	9195.8	7720.7	8564.4	8898.0	9058.5	8771.2	8908.3
Estimated effort	Mean	9.66	9.41	8.86	9.61	7.66	3.23	3.72	4.72	5.98	5.82
	SD	6.01	6.71	7.21	6.77	5.74	2.95	3.32	4.31	5.23	5.43
	Min	0.83	1.79	1.35	1.38	2.15	0.83	0.94	0.96	1.11	1.07
	Max	29.18	36.55	36.11	35.33	31.54	25.54	25.03	28.99	39.45	41.24

Table 3. Regression Model for Imputing Missing Days at Sea

Days at sea	Model 1	Model 2	Model 3	Model 4*	Model 5
Operation days	0.848***	0.815***	0.875***	0.808***	0.585***
	(0.0143)	(0.0149)	(0.0223)	(0.023)	(0.0386)
Fuel expenditure	No	4.322***	No	4.351***	2.928
		(0.641)		(0.646)	(2.141)
Constant	No	No	-3.555	2.678	67.60***
			(7.500)	(7.389)	(11.56)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Vessel fixed effects	No	No	No	No	Yes
R^2	—	—	0.624	0.641	0.505
N	964	964	964	964	964

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. We used model 4 to impute missing days at sea in the analysis. The R-squared of Model 5 is the within value from running OLS on the de-meaning data. The between and overall R-squared are 0.597 and 0.613.

Table 4. Regression Model of Effort Function

Total catch quantity	Log-Level	Log-Log
Person-years	0.090***	0.156**
	(0.02)	(0.057)
Fuel expenditure	0.039**	0.133**
	(0.016)	(0.031)
Labor compensation	0.032**	0.703**
	(0.003)	(0.041)
Constant	11.32***	10.51***
	(0.084)	(0.103)
Year fixed effects	Yes	Yes
Vessel fixed effects	Yes	Yes
R^2	0.27	0.41
N	1092	1092

Standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1

Table 5A. Rejection Rates ─ Operating days

Years / Number of Vessels	3	6	8	10
5	0.01	0.00	0.04	0.22
10	0.04	0.03	0.30	0.53
50	0.40	0.58	0.96	1.00
100	0.81	0.88	1.00	1.00
150	0.93	1.00	1.00	1.00

Table 5B. Rejection Rates ─ Imputed Days at Sea

Years / Number of Vessels	3	6	8	10
5	0.00	0.02	0.15	0.21
10	0.01	0.02	0.28	0.55
50	0.37	0.54	1.00	1.00
100	0.65	0.90	1.00	1.00
150	0.88	0.98	1.00	1.00

Table 5C. Rejection Rates ─ Length Times Imputed Days at Sea

Years / Number of Vessels	3	6	8	10
5	0.01	0.00	0.07	0.18
10	0.01	0.03	0.35	0.68
50	0.29	0.62	1.00	1.00
100	0.69	0.87	1.00	1.00
150	0.95	0.99	1.00	1.00

Table 5D. Rejection Rates ─ Estimated Total Effort

Years / Number of Vessels	3	6	8	10
5	0.01	0.00	0.09	0.19
10	0.01	0.02	0.24	0.35
50	0.22	0.49	0.98	1.00
100	0.57	0.80	1.00	1.00
150	0.77	0.95	1.00	1.00

Table 6A. Rejection Rates ─ Operating Days, with Sampling Error

Years / Number of Vessels	3	6	8	10
5	0.00	0.15	0.26	0.21
10	0.00	0.13	0.25	0.38
50	0.03	0.15	0.31	0.59
100	0.10	0.21	0.36	0.69
150	0.18	0.28	0.40	0.75

Table 6B. Rejection Rates ─ Imputed Days at Sea, with Sampling Error

Years / Number of Vessels	3	6	8	10
5	0.00	0.01	0.15	0.20
10	0.00	0.00	0.25	0.51
50	0.00	0.33	0.99	1.00
100	0.00	0.70	1.00	1.00
150	0.00	0.87	1.00	1.00

Table 6C. Rejection Rates ─ Imputed Days at Sea Times Length, with Sampling Error

Years / Number of Vessels	3	6	8	10
5	0.01	0.00	0.07	0.18
10	0.01	0.03	0.35	0.68
50	0.29	0.62	1.00	1.00
100	0.69	0.87	1.00	1.00
150	0.95	0.99	1.00	1.00

Table 6D. Rejection Rates ─ Estimated Effort, with Sampling Error

Years / Number of Vessels	3	6	8	10
5	0.00	0.00	0.00	0.06
10	0.00	0.00	0.00	0.14
50	0.00	0.13	0.34	0.97
100	0.00	0.40	0.73	1.00
150	0.00	0.47	0.92	1.00

Table 7A. Rejection Rates per Group with Weighted Sampling – Operating Days

Years	Vessels	Big-After	Big-Before	Small-after	Small-before	Diff-in-Diff
3	5	0.04	0.15	0.01	0.07	-0.05
3	10	0.28	0.30	0.08	0.23	0.13
3	50	0.92	1.00	0.57	0.97	0.32
4	5	0.19	0.16	0.05	0.08	0.06
4	10	0.53	0.40	0.16	0.30	0.27
4	50	1.00	0.99	0.89	1.00	0.12
5	5	0.16	0.10	0.05	0.09	0.10
5	10	0.48	0.46	0.18	0.29	0.13
5	50	1.00	1.00	0.90	0.99	0.09

Table 7B. Rejection Rates per Group with Weighted Sampling – Imputed Days at Sea

Years	Vessels	Big-After	Big-Before	Small-after	Small-before	Diff-in-Diff
3	5	0.13	0.07	0.04	0.09	0.11
3	10	0.34	0.27	0.15	0.18	0.10
3	50	1.00	1.00	0.80	1.00	0.20
4	5	0.26	0.11	0.04	0.05	0.16
4	10	0.48	0.28	0.24	0.21	0.17
4	50	1.00	0.99	0.96	0.99	0.04
5	5	0.23	0.14	0.11	0.07	0.05
5	10	0.57	0.40	0.27	0.22	0.12
5	50	1.00	1.00	0.99	0.99	0.00

Table 7C. Rejection Rates per Group with Weighted Sampling – Length times Days at Sea

Years	Vessels	Big-After	Big-Before	Small-after	Small-before	Diff-in-Diff
3	5	0.04	0.05	0.01	0.07	0.05
3	10	0.36	0.23	0.09	0.24	0.28
3	50	0.99	0.98	0.74	0.97	0.24
4	5	0.04	0.09	0.07	0.11	-0.01
4	10	0.40	0.34	0.18	0.26	0.14
4	50	1.00	1.00	0.82	0.99	0.17
5	5	0.20	0.11	0.07	0.08	0.10
5	10	0.60	0.33	0.19	0.18	0.26
5	50	1.00	0.99	0.95	0.97	0.03

Table 7D. Rejection Rates per Group with Weighted Sampling – Estimated Total Effort

Years	Vessels	Big-After	Big-Before	Small-after	Small-before	Diff-in-Diff
3	5	0.06	0.10	0.00	0.02	-0.02
3	10	0.21	0.17	0.03	0.08	0.09
3	50	0.98	0.95	0.61	0.79	0.21
4	5	0.09	0.12	0.04	0.05	-0.02
4	10	0.35	0.36	0.10	0.08	-0.03
4	50	1.00	0.99	0.76	0.75	0.00
5	5	0.18	0.10	0.01	0.03	0.10
5	10	0.47	0.30	0.12	0.10	0.15
5	50	1.00	0.99	0.77	0.85	0.09

Table 8. Total Number of Vessels per Group per Year

Year	Number of Vessels Per Group Per Year			
1998	Big-Before	277	Small-Before	917
1999		240		903
2000		230		851
2001		226		838
2002		253		977
Avg		245		897
2003	Big-After	263	Small-After	1178
2004		231		1111
2005		210		921
2006		197		968
2007		197		1093
Avg		220		1054

Table 9 Distance to the Model, by Vessel Size and Property Rights Regime

Measure of Effort		Small Before	Small After	Large Before	Large After	Combined
Operating Days	MSE	0.00177	0.11852	0.00540	0.09391	0.14122
	Adjusted MSE	0.19987	1.30189	0.29854	3.82339	1.44722
	KS p-val	0.87	0.00	0.63	0.17	0.00
Imputed Days at Sea	MSE	0.00038	0.20657	0.04041	0.17119	0.33418
	Adjusted MSE	0.04259	2.26897	2.23303	6.96962	3.42457
	KS p-val	0.99	0.00	0.72	0.42	0.00
Days x Length	MSE	0	0.001415	0.000348	0.001605	0.002743
	Adjusted MSE	0	0.015547	0.019244	0.065333	0.028109
	KS p-val	1.00	0.00	0.72	0.35	0.00
Estimated Total Effort	MSE	0.93228	32.90023	7.79900	22.24900	48.64777
	Adjusted MSE	105.41	361.38	430.94	905.80	498.53
	KS p-val	1.00	0.02	0.72	0.98	0.00

This table shows, for each of the four measures of effort, the mean-squared error (ie mean of the squared distances between model-consistent marginal costs and the revealed marginal costs), the mean-squared error adjusted for the number of constraints in the quadratic program (rather than the number of cells), and p-value for the KS test. Results are shown separately for large and small vessels, for before and after the property rights reform, as well as for the combined model.

Appendix A

the following statements on a panel data set $\mathcal{O} = \left\{p_t F_t, \left(q_{i,t}\right)_{i\epsilon 1\ldots N}\right\}_{t\epsilon 1\ldots T}$ are equivalent:

(A) The set $\mathcal{O}$ is consistent with the tragedy of the commons with concave production function and convex cost function.

(B) There exists a set of nonnegative numbers $\{C'_{i,t}\}_{i\epsilon 1\ldots N}$ that satisfy the linear program:

(i) $\frac{p_t \mathrm{F}_t(Q_t) - Q_t C'_{i,t}}{q_{i,t}} = \frac{p_t \mathrm{F}_t(Q_t) - Q_t C'_{j,t}}{q_{j,t}} \geq 0 \;\; \forall\, i, j \in I, \forall\, t \in T;$

(ii) $\left(q_{i,t} - q_{i,t'}\right)\left(C'_{i,t} - C'_{i,t'}\right) \geq 0 \;\; \forall\, i \in I, \forall\, t, t' \in T;$

(iii) $C'_{i,t} \geq 0 \;\; \forall\, i \in I, \forall\, t \in T.$

Proof

Our proof is straightforward and follows the outline of Carvajal et al. (2013). To see (A) implies (B), suppose that the data are rationalized with production $\{p_t F_t, q_{i,t}\}_{i\epsilon 1\ldots N, t\epsilon 1\ldots T}$. Then the first order condition guarantees the existence of $\{C'_{i,t}\}_{i\epsilon 1\ldots N}$ that satisfy the common ratio property (i). Given convexity of costs, the co-monotone property (ii) is satisfied as well.

To see (B) implies (A), we first show that at observation t, when (i) is satisfied, there exists a concave production function F_t such that $\bar{F}_t(Q_t) = F_t$, and with each firm having the cost function $\bar{C}_i$, $\{q_{i,t}\}_{i\epsilon 1\ldots N, t\epsilon 1\ldots T}$, which constitutes behavior consistent with the Tragedyofthe–Commons model. We define $\bar{F}_t(Q_t)$ by $p_t \bar{F}'_t(Q_t) = \frac{p_t \bar{F}_t(Q_t)}{Q_t} - b_t$ and let $b_t = \frac{p_t \mathrm{F}_t(Q_t) - Q_t C'_{i,t}}{q_{i,t}}$. A concave function will satisfy the definition here since the average return is larger than the marginal return. Firm i's decision is to choose $q_{i,t}$ that maximizes profit $\left\{\frac{q_{i,t}}{Q_t} * p_t F_t(Q_t)\right\} - C'_{i,t}$; this function is concave, so the input level is optimal if and only if it obeys the first-order condition. Apply $\bar{F}_t(Q_t)$ defined above, we have $\frac{q_{i,t}}{Q_t} * p_t \bar{F}'_t(Q_t) + \left(1 - \frac{q_{i,t}}{Q_t}\right) * \frac{p_t \bar{F}_t(Q_t)}{Q_t} - C'_{i,t} = \frac{q_{i,t}}{Q_t}\left(\frac{p_t \bar{F}_t(Q_t)}{Q_t} - \frac{p_t \bar{F}_t(Q_t) - Q_t C'_{i,t}}{q_{i,t}}\right) + \left(1 - \frac{q_{i,t}}{Q_t}\right) * \frac{p_t \bar{F}_t(Q_t)}{Q_t} - C'_{i,t} = 0$. Hence, $q_{i,t}$ is the profit-maximizing input of firm i at time t.

Second, we show that if for some firm i there are positive scalars $\{C'_{i,t}\}_{T\epsilon 1\ldots T}$ that are increasing with $q_{i,t}$, then there exists a convex cost function $\bar{C}_i$ such that $C'_{i,t} \epsilon \bar{C}_i(q_{i,t})$. Proof of this part is the same as in Lemma 2 in Carvajal et al. (2013).

Using the two conclusions above, we see that constraint (i) confirms that the choice of input $q_{i,t}$ is the optimal choice that satisfies the first order condition of the TOC model. And constraints (i) and (ii) ensure that marginal costs revealed from the linear program is the taken from a time-invariant convex cost function. Constraint (iii) ensures the nonnegativity of marginal costs. Hence, satisfying the three properties in the linear program implies consistency with the TOC model.