6-11-2019

Essays on the Commons, Scarcity and Innovation

Yaqin Liu

ESSAYS ON THE COMMONS, SCARCITY AND INNOVATION

by

YAQIN LIU

Under the Direction of H. Spencer Banzhaf, PhD

ABSTRACT

This dissertation contains three essays. In the first essay, I derive non-parametric tests of behavior consistent with the tragedy of the commons model based on recent results in the industrial organization literature (Carvajal et al. 2013). The 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. I extend the tests to account for sampling 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. Applying this approach to panel data of Norwegian fishers, I find evidence rejecting the tragedy of the commons model. Significantly, I find that rejection rates of the model increase after property rights reforms moved the fishery away from the tragedy of the commons.

In the second essays, I bridge the research in behavioral economics and environmental psychology and use designed experiments to test the effects of natural and urban environments on risk and time preferences. I examine whether time in green /urban space can improve economic decision making, helping people escape the poverty trap. Results from the laboratory show evidence that viewing urban pictures tend to reduce cognitive burdens and decrease risk averse behavior among people familiar with urban environment.

In the third essay, I examine the effects of a carbon tax on energy-efficient innovations. In an influential paper, Popp (2002) empirically analyzed the effect of energy prices on energy-efficient innovations as measured by new patents. In this research, I explore the efficiency of carbon taxes, or the extent to which they can induce innovation through the price channel. Based on the estimates from Popp's model, I estimate the costs directly added from a carbon tax based on emission coefficients of different types of fuels in the U.S. market and predict the number of new energy-efficient innovations that would be stimulated by a carbon tax. Results from the estimation indicate the significance of the price elasticity of energy-efficient innovations and limited effects from scaling up carbon taxes.

INDEX WORDS: Tragedy of the commons, Nonparametric test, Cognitive burden, Nature, Carbon tax, Technology Innovation

ESSAYS ON THE COMMONS, SCARCITY AND INNOVATION

by

YAQIN LIU

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2019

ESSAYS ON THE COMMONS, SCARCITY AND INNOVATION

by

YAQIN LIU

Committee Chair: H.Spencer Banzhaf

Committee: Martin D. Smith

Garth Heutel

Mike Price

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 NON-PARAMETRIC TESTS OF THE TRAGEDY OF THE COMMONS

1.1 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 decline in marginal productivity 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, extracting oil from a common pool, and fishing from the sea.[1]

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. In the context of fisheries, Costello, Gaines, and Lynham (2008) and Birkenbach, Kaczan and Smith (2017) find that individual catch shares can prevent the collapse of fisheries and slow the race to fish. This policy outcome is consistent with over-exploitation in the open-access

[1] Compare also Sen's (1966) model of cooperatives.

regime, but does not test individual behavior. Kirkley, Paul, and Squires (2002) and Felthoven, Horrace, and Schnier (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.

On the other hand, Huang and Smith (2014) have 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 vessel 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. Our approach tests whether individual behavior is consistent with the tragedy of the commons model. In particular, drawing on recent work by Carvajal el al. (2013), we develop a non-parametric revealed preference-type test for the "tragedy" in common pool resources. Carvajal et al. developed a revealed preference test for Cournot equilibrium, deriving properties that hold when firms are strategically interacting as predicted by the Cournot model. As the tragedy of the commons and the Cournot model are essentially isomorphic (both are surplus-sharing games), we can derive similar properties that hold under the strategic interactions of the tragedy of the commons. This approach 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.

We extend the 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 analysis 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 rejections 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

Diewert (1973) and Varian (1985), we include an adjustment factor in the model to guarantee that data would always pass the model. We apply a linear program to reveal the minimal magnitude of the adjustments required. Based on this result, we apply a chi-squared test and a Kolmogorov-Smirnov test to inform probability distributions for rejections of the model. Variations of these extensions could also 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). As a quantitative measure of the extent of rejection, we subsample from the data set and obtain rejection rates from repeated subsampling. 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. Lastly, the statistical test based on distance from revealed marginal costs to model-consistent marginal costs also rejects the hypothesis that data is consistent with the TOC model. Interestingly, preliminary results show that the rejection rates are higher after institutional reforms in the Norwegian fishery that reduced open access.

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.

1.2 Theoretical Model: A Nonparametric Test of the Tragedy of the Commons

1.2.1 The Case with Time-Invariant Cost of Effort

Consider an industry consisting of I profit-maximizing firms, indexed by $i = 1,2, \ldots, I$, each having free access to an exogenously fixed common property resource. There are T decision periods indexed by $t = 1,2, \ldots, T$. 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 catch 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 and which—for now—we treat as time invariant.

Following Carvajal et al.'s logic for Cournot competition, we say a panel data set $\mathcal{O} = \left\{p_t F_t, \left(q_{i,t}\right)_{i\epsilon 1 \ldots N}\right\}_{t\epsilon 1 \ldots T}$ 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 as do Kirkley, Paul, and Squires (2002) or Huang and Smith (2014). 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. These assumptions are sufficient to guarantee a quasi-concave profit function to be maximized.

To see this, note that firm *i*'s profit-maximization problem at time t is:

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

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

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

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 (Gordon 1954).

Rearranging terms, we obtain:

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

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.:

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

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*,

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

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.

1.2.2 Linear Program for the Test

Our approach to test 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 agent *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 consistency with utility-maximizing behavior using the Generalized Axiom of Revealed Preference (GARP) (Afriat, 1967). This 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 (See Appendix A for a proof):

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

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 $\{C'_{i,t}\}_{\forall\, i \,\epsilon\, I, \forall\, t \,\epsilon\, T}$, will result a rejection to the model in this basic test.

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). It is always the case, of course, that failure to reject a null hypothesis does not guarantee it is true. 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 regulations, then arguably it is the auxiliary hypotheses that are being rejected, not the fundamentals of the behavioral model.

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, fisheries with individual fishing quotas (IVQs) that are non-tradable. IVQs restrict each vessel to catch up to its quota. Although non-tradability prevents cost minimization subject to a total catch (as firms with high costs at the margin may be allocated quota that cannot be traded to low-cost firms), IVQs have some advantages. Typically, they cap the total allowable catch so as to protect the sustainability of a fishery. Additionally, unlike group quotas (which also cap the total catch), they prevent a "race to fish" within a season, as a firm's share is exogenous to how quickly it fishes.

Importantly, IVQs 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:

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

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:

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

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

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

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 $C'_{i,t}/(p_t - \lambda_{i,t})$, 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.

1.3 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.

1.3.1 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 means times N.

Suppose we observe $p_t\hat{F}_t = p_t F_t * \alpha_t$ and $\hat{Q}_t = Q_t * \beta_t$, instead of the true total revenue, where α_t and β_t are proportionate errors. Then the common ratio property becomes $\frac{\alpha_t p_t \mathrm{F}_t(Q_t) - \beta_t Q_t (C'_{i,t})}{q_{i,t}} = \frac{\alpha_t \lambda_t p_t \mathrm{F}_t(Q_t) - \beta_t Q_t (C'_{j,t})}{q_{j,t}}$. Dividing both sides by β_t and letting $\lambda_t = \alpha_t/\beta_t$, we can write the linear program with sampling errors as:

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

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

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

Without sampling errors, we should look for marginal costs that satisfy properties above without λ_t. We treat λ_t as unknowns and let the linear program look for the set of $\{\lambda_t, C'_{i,t}\}_{\forall\, i \in I, \forall\, t \in T}$ 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.

In the linear program specified above, λ_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 $\{\lambda_t\}_{\forall\, i \in I, \forall\, t \in T}$ 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.

1.3.2 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. A natural question to ask is how "low" is low enough for us to attribute the rejections to nuanced situations such as sampling errors or trembling hands in a small portion of participants' behavior. On the other hand, how "big" of a rejection is big enough for us to statistically conclude the data does not conform to the model? That's the question we will answer in this section.

Building on the marginal-cost-consistency method described in 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 data to be satisfied 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. The adjustment factors are constructed in a way to guarantee that data would always pass the model. We deploy 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 a chi-squared test and Kolmogorov-Smirnov test to inform statistical acceptance/rejection of the model.

We use the following linear program:

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

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

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

$\delta_{i,t}$ is the minimum adjustment factor on marginal cost $C'_{i,t}$. Constraints (i), (ii) and (iii) guarantee that the set $\{\delta_{i,t}, C'_{i,t}\}_{\forall\, i \in I, \forall\, t \in T}$ satisfies the common-ratio property, co-monotone property, and nonnegativity constraint. By construction, such solutions always exist.[2] Hence, we

[2] This is because the adjustment factors expand the domain of marginal costs to all real numbers. As there are 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

can identify and quantify the minimal adjustment factors $\{\delta_{i,t}\}_{\forall\, i\, \in I, \forall\, t\, \in T}$, which is the minimal distance between the revealed marginal costs to the model-consistent marginal costs.

1.3.2.1 A Chi-squared Test

Taking the minimal distance found above, we can conduct a chi-squared test of the null hypothesis that the data is consistent with the model. Denote the set of marginal costs that is 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 $C'_{i,t} + \delta_{i,t}$. Assume the model-consistent marginal costs follow a log-normal distribution $N(\mu, \sigma^2)$ with the lower limit zero. Finally, 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 $C'_{i,t}$.

Under the null hypothesis that an observed data set is consistent with the model, revealed marginal costs would converge to the distribution of model-consistent marginal costs in the limit. Hence, $\mathrm{z_{i,t}}=\frac{\log(\widehat{\mathrm{mc_{i,t}}})\text{-}\log(\widetilde{\mathrm{mc_{i,t}}})}{\sigma}$ follows a standard normal distribution. And we can easily obtain $z_{i,t}$ from the adjustment factor $\delta_{i,t}$ solved from the linear program, given $\delta_{i,t} = \widehat{mc_{i,t}} - \widetilde{mc_{i,t}}$. As a result, $S = \sum_{t=1}^{T} \sum_{i=1}^{I} {z_{i,t}}^2/\sigma^2$ follows a chi-squared distribution with $T * I$ degrees of freedom. With 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.

1.3.2.2 Two-sample Kolmogorov-Smirnov Test

Alternatively, we can test the null hypothesis with the Kolmogorov-Smirnov (KS) test.

---

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.

The advantage of this method is that we do not need to assume a particular distribution of the model-consistent marginal costs. 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 contains the revealed MCs. And the sample size for both samples is $I * T$. $D_{n,m}$ is a vector consisting 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.

1.3.3 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 use the linear program:

$$\min_{c'_{i,t},\delta_{i,t}} \sum_t \sum_i |\delta_{i,t}|$$

Subject to:

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

(v) $[(q_{i,t} + \delta_{i,t}) - (q_{j,t} + \delta_{j,t})](C'_{i,t} - C'_{i,t'}) \geq 0 \quad \forall\, i, j \in I, \forall\, t \in T;$

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

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. In future work, we will consider using this approach.

1.4 Empirical Application

In this section, we describe the Norwegian fishery and the dataset to which we apply our tests of the tragedy of the commons.

1.4.1 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 whitefish 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 consists of 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 consists of smaller vessels using a variety of gears such as long lines, troll nets and Danish seine. Our sample contains only the coastal fleet, with no trawlers. Each fishing vessel is separately owned by an operator, so vessels can be taken as firms in our model.

In 1989, after the collapse of the Northeast Atlantic cod, a total allowable catch (TAC) quota was set by the joint Norwegian-Russian Fishing Committee for whitefish, with the TAC divided between the trawler fleet and the coastal fleet. Then in 1990, an IVQ system was theoretical introduced to the Norwegian coastal fleet. However, many fishermen (particularly small vessel owners) were allocated larger quotas than their previous catches, whereas others (mostly large vessels) were allocated smaller quotas than their previous catches. To ensure that the allocated quotas were fished within the coastal vessel group, an "overbooking system" was introduced in 1991, which allowed vessel owners to fish above the allocated quotas. As the overbooking was substantial, the IVQ system essentially was not binding during the early years of the program, making it 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 total allowable catch (TAC) quota 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 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. This problem appears to be especially problematic for the small vessels in 2003.

Finally, in 2004, vessels above 15 meters were allowed to merge quotas from several vessels onto one vessel. Meanwhile, overbooking was no longer allowed. 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. See Hannesson (2013), Standal, Sønvisen and Asche (2016), and Standal and Asche (2018) for further discussion of the fishery.

In sum, from 1998 to 2002, all vessels were under an RRA regime. After 2003, big 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.)

1.4.2 Description of Data

The data for the Norway coastal fleet come from an annual random survey of vessels from 1998 to 2007. Only a sample of the registered active vessels are surveyed each year. The first row of Table 1 shows the sample size (number of observations). The second row of Table 1 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. Additionally, based on surveys of all fishers, we observe annual data 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 revenues received by species (cod, haddock, saithe and other whitefish). We also have catches of cod, haddock, saithe and other whitefish, both in physical units (tons) and revenues (NOK). 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. It shows a clear upward trend in total revenue, with each year higher than the previous. Row 4 similarly shows total catch in tons, which follows a similar upward trend. The remainder of Table 1 offers additional details on the distribution of catch across vessels and across species, by year.

Although it requires only annual aggregate revenue on the output side, our test requires micro-level data on the input side. Vessels do not necessarily fish in all years, 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 2 shows raw data on inputs, including operating days, days at sea, person-years, labor compensation, and fuel expenditure.

1.4.3 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 3 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 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 4 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 3 shows annual data on imputed days at sea.

Our third measure of input uses these imputed days at sea times vessel length. This is a common proxy for input in the fishery literature. Table 3 also reports annual values of this product.

Our fourth and final measure of input aggregates multiple input variables into a scalar-valued function. This 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

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

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

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

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. Results of estimating this model are shown in Table 5. 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 3 shows summary statistics for this estimated value.

1.4.4 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 $T \in \{3, 6, 8, 10\}$.[3] 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.)

1.4.5 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) by before (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

[3] The larger the number of vessels and windows, the longer it takes to run the test. Due to resource constraints, the largest number of vessels we sample is 150. The subset with 150 vessels and 13 years takes 16 days to run on an i7-4770 CPU, 24GB, 64-bit computer.

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 probabilities of being selected among vessels of later periods (2003-2007). Probabilities of 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 up 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, hence 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. It helps to balance the number of observations among groups to generate credible difference-in-difference results. 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-difference to inform effects of regime changes.

As discussed in Section 2.2, data generated from the IVQ regime is not expected to be consistent with the tragedy of the commons model. We expect the difference of big vessels (after-before) will be higher than those for small vessels.

1.5 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).

1.5.1 Results of Test Pooling all Data

Tables 6-9 present results using the basic test of Section 2, using four respective proxies of effort: operating days, imputed days at sea, length times days 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. Note that the rejection rates generally are increasing in N (moving down the rows) and T (moving to the right across columns), as the number of equations and inequalities to satisfy is increasing in these parameters. Exceptions to this rule are due to random sampling. Furthermore, when more than 100 vessels are considered for longer than 6 years, the rejection rates approach one. These results indicate that the behavior of vessels/fishermen in our sample cannot be explained by the TOC model when a fair 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 a large number of missing observations in the data set[4]. 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

[4] 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.

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. Tables 10-13 present results using the test with sampling errors. 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.

1.5.2 Results of Statistical Tests

We conduct the Chi-squared test of Section 3.2 to the same subsamples in the basic tests for input operating days. For now, we only take subsamples for $N \in \{5, 10, 50\}$ and $T \in \{3, 6\}$. The test statistics of operating days is $T_{oper} = \sum_{t=1}^{T} \sum_{i=1}^{N} {z_{i,t}}^2/\sigma^2 = \frac{57286}{0.3684} = 12365.8$. The chi-squared critical value $C_{0.05,58500}$ = 59064. Thus, we reject the null hypothesis that the data obtained for a group of vessels with input operating days is consistent with the behavioral model depicted in the tragedy of the commons. Chi-squared tests of other three input variables (imputed days at sea, length times days and estimated effort) yield similar results.

We also conduct the KS test for the same subsamples of operating days as in the Chi-squared test. We first plot the distributions of the two samples of marginal costs in Figure 1. In the figure, the revealed marginal costs are $\{\widehat{mc_{\iota,t}}\}_{i \in I, \forall\, t \in T}$ and the model-consistent marginal costs are $\{\widetilde{mc_{\iota,t}}\}_{i \in I, \forall\, t \in T}$. From Figure 1, we can narrow down the maximum distance of the two distributions to the intervals of marginal costs between [1, 2]. To find the maximum distance of the CDFs between the samples, we take bins of 0.1 within the interval and compare the two CDFs. Results are presented in Table 14. The estimated maximum distance $D_{N*T,N*T}$= 0.18, and the critical value is 0.019. Hence, again, we reject the null hypothesis that the data obtained for a group of vessels with input operating days is consistent with the model. KS tests of the other three inputs (imputed days at sea, length times days and estimated effort) yield the same conclusion.

Results from the two statistical tests confirm our observation from the rejection rates tables in Section 6.1. More importantly, the maximum distance obtained from the KS test can be used as an indicator to inform how far the current regime in the Norwegian ground fishery is away from the TOC model. Consider one fishery, we can track the maximum distance overtime and evaluate the effectiveness of management policy at different time periods. Across fisheries, the method can be used to gauge the distance to the TOC setting for fishing behavior under different property-rights management regimes. Hence, it can be adopted to compare efficiency levels across sites. Since the test does not rely on parametric specification of production, comparison across sites do not need to estimate production or cost function for each site.

1.5.3 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 divided 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 15-18 present results per group after weighted sampling. 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 weighted sampling as described in section 4.5 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 are reflecting 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 moved closer to Nash in 2003. One possible explanation for this finding is an induced race to fish after securing the small group right without assigning rights to individuals. Comparing the number of small and big vessel across years (see Table 19), there is a marked increase in the total number of small vessels in each year starting in 2003, whereas there is not much change in 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.[5] After 2003, separated TAC for small group reduced the competition from big vessels and secured a potential economic rent. However, without individually assigned property rights to quotas, the secured economic rent attracted new entrants and spurred the race to fish. This result 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 speaks to the findings in Kroetz et al. (2015) that policy with good social objectives could reduce overall efficiency and rents in fisheries.

1.6 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) using 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 groundfishery. Overall, we find the behavior of individual fisherman/vessel of the Norwegian Coastal Fishery does not conform to the model of

[5] 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.

the tragedy of the commons. However, we find that rejection rates are larger after property rights reforms, for the large vessels that received stronger property rights. Our approach can be applied to other common pool resources whenever individual effort data is available. And it can be used to compare effectiveness of different management regimes. It yields important policy implications. For instance, despite their theoretical appeal and mounting empirical evidence of economic benefits, property rights-based management in fisheries remains controversial. Critics of catch shares express concerns about social issues such as equity and effects on fishing communities. These views are expressed in the policy process that ultimately shapes how fisheries are regulated. Do the resulting regulations address the first-order problem of the commons, or do they preserve the tragedy? Our approach can help to answer this question and suggest which regulatory regimes are not up to the task. After all, even the strongest opponents of catch shares do not defend the tragedy of the commons as an appropriate alternative.

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
Total annual value (100 million NOK)		3.61	3.67	3.67	3.91	3.98	4.54	4.61	4.68	6.58	7.40
Total 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 2 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 3 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	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	190.0	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 4 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 demeaning data. The between and overall R-squared are 0.597 and 0.613.

Table 5 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 6 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 7 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 8 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 9 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 10 Rejection Rates — Operating Days with Measurement 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 11 Rejection Rates — Imputed Days at Sea with Measurement 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 12 Rejection Rates — Imputed Days Times Length with Measurement 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 13 Rejection Rates — Estimated Effort with Measurement 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 14 CDF of Two Samples

MC	**CDF_**$\widehat{mc_{\iota,t}}$	**CDF_**$\widetilde{mc_{\iota,t}}$.	**Distance**
1.0	0.60	0.44	0.16
1.1	0.65	0.48	0.17
1.2	0.70	0.52	0.18
1.3	0.75	0.56	**0.18**
1.4	0.79	0.61	0.18
1.5	0.83	0.64	0.18
1.6	0.86	0.68	0.17
1.7	0.88	0.71	0.17
1.8	0.91	0.75	0.16
1.9	0.92	0.78	0.15
2.0	0.94	0.80	0.14

Table 15 Rejection Rates per Group - 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 16 Rejection Rates per Group – 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 17 Rejection Rates per Group – 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 18 Rejection Rates per Group – 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 19 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
Ave		**245**		**897**
2003	Big-After	263	Small-After	1178
2004		231		1111
2005		210		921
2006		197		968
2007		197		1093
Ave		**220**		**1054**

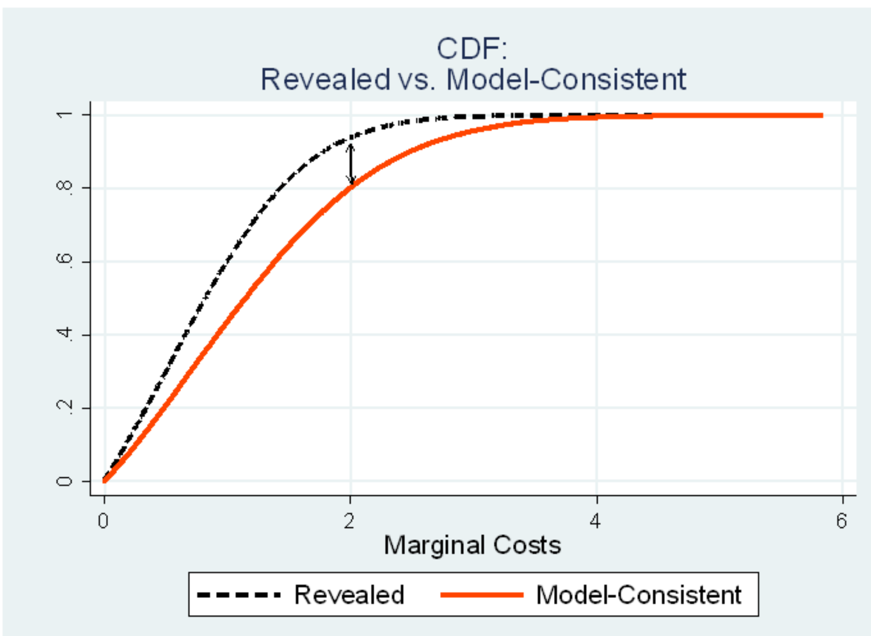

Figure 1 Distribution of Two Samples

2 NATURE AND SCARCITY

2.1 Introduction

Research at the intersection of economics and psychology has found that cognitive burdens hinder people's ability to perform well economically and to make prudent economic decisions. They perform worse at tasks requiring attention, become more cautious, and become more impatient (Deck and Jahedi 2015). In particular, the cognitive burdens associated with acute scarcity of time or money appear to create tunnel vision in which people focus on pressing needs while neglecting long-term consequences (Shah et al. 2012, Mani et al. 2013, Mullainathan and Shafir 2013, Haushofer and Fehr 2014). It appears that when "System 2," or analytical thinking, is burdened by urgent present needs, less cognitive capacity is available for use in other judgments and decisions (Kahneman 2011, Schilbach, Schofield, and Mullainathan 2016). This vicious cycle may perpetuate poverty traps.

At the same time, research at the intersection of psychology and health has found that the sense of beauty and wonder inspired by green environments improves one's mind, body, and spirit. The restorative rest of time in nature improves focused attention and reduces stress (Bowler et al. 2010, Bratman et al. 2012, 2015, Hartig et al. 2014). Relative to time spent in brown, urban environments, time in green, natural environments has been found experimentally to improve short-term recall and performance at tasks requiring attention (Berman et al. 2008, Berto 2005, Hartig et al. 2003). It also reduces stress, as seen in lower blood pressure, salivary cortisol, and neural activity in the subgenual prefrontal cortex (associated with rumination), and to improve self-reported mood. The effects extend to looking out windows to or at pictures of green spaces (Berman et al. 2008, Berto 2005, Hartig et al. 1999, 2003, Hartig and Staats 2006, Ulrich et al. 1991). These experimental findings also appear to have external validity. In observational studies, individuals living closer to green spaces had lower cortisol levels and lower mental distress (Thompson et al. 2012, White et al. 2013). Additionally, children who moved homes show significant cognitive functioning improvement among those who move to greener areas (Wells 2000).

In this research, we explore the intriguing possibility that, as they facilitate focused attention and reduce stress, green environments concomitantly improve economic decision making. We test this hypothesis by conducting lab experiments replicating previous work on cognitive burdens and economic decision making (Deck and Jahedi 2015, Shah et al. 2012), but crossing them with time spent viewing pictures of green vs. urban scenes, as in many environmental psychology experiments (e.g. Hartig and Staats 2006). We test whether exposure to green/urban environments affects subjects' performance on tasks requiring attention, their risk aversion, and their patience. We also test whether it especially improves such performance among participants whose focused attention has been depleted by greater cognitive burdens.

Results from experimental data with undergraduate college students do not show any effect of nature on restoring cognitive capacity. On the contrary, there is evidence that viewing pictures of urban scenes helps to reduce risk averse behavior among subjects who are burdened by memorizing long digit numbers. Below we discuss the details of our experimental design and issues that need to be addressed in future research.

2.2 Experiment

A total of 246 undergraduate students from Georgia State University (GSU) were recruited for this study. The experiment was clustered in sessions of approximately 20 people at the Experimental Economics Center (ExCEN) of Georgia State University. The participation fee was $10, and participants were paid an additional $15 on average. All participants were seated at a computer before the experiment started. Hard copies of experimental instructions were provided to each subject. The experimenter read the instructions to participants and guided practice rounds prior to the start of the experiment. Practice rounds were designed with feedback to make sure that participants understood their choices and associated payoffs from their choices.

2.2.1 Methodology

Drawing mainly on the design in Deck and Jahedi (2015) but also work in psychology (e.g.Berman et al. 2008), in the experiment we use a digit-memorizing task to induce varying degrees of cognitive load. Half the individual participants in each session were randomly assigned to memorize a 1-digit number, the other half to memorize a 5- to 8-digit number. Considering subjects may give up memorizing if correct recall become too difficult, we modified the long 8-digit number in Deck and Jahedi to numbers varying between [5, 6,7,8] digits, with subjects started with 5-digit and increased gradually to up to 8-digit given their correct recall in previous periods, likewise, the number decreased down to 5 if their previous recall is incorrect. Hence, the length of number was different for each participant and solely depended on participant's performance in memorization during the experiment.

The experiment contains 80 repeated periods. At the beginning of each period, a randomly selected number of the predetermined length appeared on their screen for five seconds. They were instructed to memorize that number without making notes. The payoff for correctly recalling the number was $22. And $0 would be paid if the number recalled was wrong. The payoff for correct recall was higher than the average return from all other tasks in this experiment to motive subjects, to effectively induce cognitive load.

Next, subjects were presented with one question randomly drawn from one of the three main tasks, which are explained in more detail below: (1) a dots-mixed task, (2) a gamble for small stakes, or (3) an inter-temporal choice. Once the question was completed, participants were asked to recall the number they just memorized and enter it in a blank box on the screen. The next screen told the participants if their answers were correct or not. Then the second period began, a new digit would appear, and participant were presented with another question, again randomly drawn from the three tasks.

In the Dots-Mixed task, participants saw pictures of either a green apple or a lime presented on the left or right side of a fixation cross. If they see a green apple, they should press a key on the same side as the apple ('Q' for left, 'P' for right). If they see a lime, they should press a key on the opposite side ('P' for left, 'Q' for right). The order of the trials was randomized. Each trial started with a fixation cross presented for half of one second. Then the apple or lime was presented for three-quarters of a second, during which participants needed to respond. The payoff for correctly pressing the key was $10. We first implemented this task as subjects use mouse to choose letters P or Q, later (starting in session 7) we modified it to using keyboard. In data analysis below, we treated analysis before and after session 7 for this task separately.

In the risk-aversion task, participants chose between two options: one that returned an amount of money $Y with certainty, or a lottery that had a 50/50 chance of returning either $0 or a larger amount of money $Z. The framing was either in terms of gains or losses. For example, one could choose between receiving $8 for sure vs. a 50/50 gamble between winning either $0 or $18; or alternatively one could be given an endowment of $16 and face a choice of losing $8 for sure or facing a gamble of losing either $0 or $14. In gains gamble, Y and Z were randomly drawn with equal probability from the pairwise set {8,18; 9, 20; 10, 22; 11, 24; 12, 26; 13, 28; 14, 30; 15, 32}. In losses gamble, endowment, Y and Z were randomly drawn with equally probability from the set {16, -8, -14; 18, -9, -16; 20, -10, -18; 22, -11, -20; 24, -12, -22; 26, -13, -24; 28, -14, -26;

30, -15, -28}. These gambles were designed to secure the same mean return in the losses and gains frames in order and fix the difference in mean return as payoff increases. In such a way, it is convenient to compare the number of safe choices (certainty amount) within subjects by counts, and to compare the change in number of safe choices across treatments.

Finally, in the time-preference task, participants were asked which they prefer: $15 of money today or $X a week from today. X was drawn with equal probability from the set {$15.25, $15.50, $15.75, $16, $17, $18, $19, $20}. To reduce the "transaction cost" associated with postponed payments, as well as to increase confidence that future payments will arrive, we adapted the procedure of Andreoni and Sprenger (2012). Participants were paid $10 in cash for completing the experiment and their payoffs from the time-preference task were mailed “today” or “in a week” regardless of their choices. Given that our sample pool mainly live on campus or close by, the time they received mailed payments were similar and were consistent with their choices.

After 40 periods, participants viewed either 50 pictures of green scenery or of urban scenery, assigned randomly. Specifically, half of the participants who memorized 1-digit numbers viewed green pictures, and the other half viewed urban pictures; likewise for those who memorized 5- to 8-digit numbers. Viewing pictures is an established way to mimic time in green or urban places and has similar effects in psychological studies (e.g. Berman et al. 2008, Berto, 2005, Hartig et al. 1999). The pictures used were the same as those used in previous work (Hartig and Staats 2006, Berman et al. 2008). Each picture was displayed for 10 seconds on the screen. Picture-viewing lasted for 10-12 minutes. At the end of picture viewing, participants were asked to rate on a scale of 1 to 5 how much they like the set of pictures (1=dislike, 5=like a lot). Participants were not paid for viewing pictures.

After viewing pictures, participants repeated the same procedure as before the pictures for another 40 periods. Those who memorized a 1-digit number before viewing the pictures continued with memorizing a 1-digit number, and those who memorized a longer digit number also continued to do so. At the end of the experiment, the computer randomly chose one of the 80 periods for which the participant gets paid. It was randomly determined whether the participant was paid by the question task or the memorization task in the chosen period.

This experiment essentially follows the design of Deck and Jahedi (2015), with some minor modifications (such as using real rather than hypothetical inter-temporal decisions, and replacing mathematical multiplication task with Dots-mixed task), and with the main modification of taking time to view pictures at the midway point. In addition, we modified the long 8-digit number in Deck and Jahedi to numbers varying between [5, 6,7,8] digits, considering subjects may give up memorizing if correct recall become too difficult. Deck and Jahedi found that participants burdened with memorizing 8-digit numbers performed more poorly on the mathematical multiplication task, showed less tolerance for risk, relative to those assigned a 1-digit number.

2.3 Results

Table 20 shows summary statistics of choices before viewing pictures by treatments of short and long digits. We can compare the results in Table 20 to summary statistics reported in Table 3 of Deck and Jahedi (2015) (hencehorth, D&J). Results for short-digit numbers are similar in our case from D&J. Given that our long-digit numbers are shorter than their 8-digit numbers on average, the burdens induced in our case should be lower. This is indicated by the percentage of correct memorization in the long-digit case in comparison to their correction rates in the multiplication task (69.5% vs. 43.3% in D&J). In our case, there is no evidence that response in the Dots-mixed task was affected by the burden of memorizing. However, D&J find multiplication task was negatively affected by memorization (71.6% correct response in 1-digit and 55.9% in 8-digit). Moreover, D&J find less risky choice (more risk aversion) in the 8-digit scenario for both gains and losses frames. In our case, similarly, there is more risk averse/safe choices in the gains lottery under long-digit numbers, but no evidence of risk aversion in the losses lottery. Lastly, D&J show more patience in time choices under 8-digit numbers, while we find less patience in the long-digit case.

In D&J, 8-digit numbers decreased the correct response in the multiplication task from 71.6% for 1-digit numbers to 55.9%. However, in our Dots-mixed task (which replaced the multiplication task), the effect of long-digit numbers is relatively minor. Also, our correct memorization in long-digit numbers is much higher than in D&J (69.5% vs. 43.3% in D&J). Our initial concern that 8-digit numbers may be too difficult, and that subjects may give up memorizing might not have been accurate. Additionally, reduced difficulty may have weakened the load and further diluted restorative effects of picture viewing treatments that might otherwise be salient under heavier cognitive burden.

Next, we show results of statistical tests on the effects of pictures and digits. First, Table 21 shows statistical tests on the effect of the cognitive burden from long-digit numbers, taking the short-digit numbers as a baseline. Since we modified the way subjects input their responses in the Dots-mixed task at session 7, we run the test for sessions post 7 and for all pooled sessions separately for the Dots task. There is weak evidence that the digit load decreased the correct response rate in the Dots task after the modification. There also is weak evidence that digit load affects risk preference in the gains context. The results raise the possibility that our reduced load might not be strong enough to change behavior. And it is unclear why D&J find less choice of the early option (more patience) in time choices, while we find less patience with long-digit numbers.

Second, we show the average treatment effects of nature pictures, taking urban pictures as the control group. Table 22 presents results when pooling different digit-number groups. It shows the difference-in-differences, that is, the difference after viewing nature pictures relative to before, vs. the same difference after viewing urban pictures. Thus, it nets out any effect from viewing pictures generally, or of taking a break from the tasks. The results show some evidence of the average treatment effects on risk choices, but otherwise little effect of nature pictures. Thus, we are unable to replicate the effect of nature viewing, using only pictures, previously reported in the literature. It may be that picture viewing is not a substitute for time spent in nature. It may also be that nature viewing does not affect these particular tasks.

Third, we turn to statistical tests of the restorative effects of nature/urban pictures. We present summary statistics for before and after picture-viewing treatments, for short and long-digit numbers respectively in Table 23. On testing the restorative effects of pictures, we are looking for changes of responses in three tasks after picture viewing compared to before picture viewing. To account for the order effects and learning effects, which include the fact that subjects can perform differently just because they take a short break during picture viewing or because they have developed skills in the tasks during the first 40 periods, we take difference-in-differences before and after picture viewing between nature and urban pictures. Thus, we take the triple difference comparing the difference-in-differences from before/after and nature/urban with long digits to the difference-in-differences with short digits. Results are presented in Table 24.

In Table 24, we run OLS regression with robust standard errors of each task response on treatments and interacting terms. Within the Nature and Urban picture scenarios, the coefficients on the interacting term After (picture viewing) times Long (digit numbers) measure the effect of picture viewing on task response of treatment group (long-digit), taking the effect on the control group (short-digit) as benchmark. Denote y as outcome, n for nature and u for urban; l for long-digit and s for short-digit numbers; a for after and b for before picture viewing. The coefficients on After×long measure the mean difference measured by: $(y_{ula} - y_{ulb}) - (y_{usa} - y_{usb})$ within the urban picture scenario, and $(y_{nla} - y_{nlb}) - (y_{nsa} - y_{nsb})$ within nature picture scenario, respectively. The triple difference of the restorative effects of nature pictures over urban pictures is measured as $[(y_{nla} - y_{nlb}) - (y_{nsa} - y_{nsb})] - [(y_{ula} - y_{ulb}) - (y_{usa} - y_{usb})]$. And the mean difference of this triple difference is captured by the coefficients on the interacting term Nature× long×After.

The bottom row of the last three columns in table 24 shows the triple-differences. We find an increase in safe choices from green pictures relative to urban pictures (for long digits relative to short digits). We find no evidence of an effect in the other tasks. Looking at the difference-in-differences, we see that urban pictures reduced risk safe choice and there is no effect from nature pictures. Thus, the significant effects on risk choices in triple difference is mainly driven by urban pictures. This finding contradicts our hypothesis that nature pictures reduce cognitive burden while urban pictures do not. As cognitive burden leads to risk averse behavior, if nature pictures effectively restore the mind and urban pictures do not, we shall see less risk safe choices with nature and more risk safe choice with urban. Results in table 24 raise two possibilities. One possibility is sample selection. Our sample are all undergraduate students from Georgia State University (GSU) and most of them have lived and are living in urban environment. In our sample, 57.3% of the subjects live on campus (in the downtown Atlanta area) and most of them have lived close by the city of Atlanta in the past. Although urban pictures were rated lower than nature pictures (see table 23), it may be that the urban settings were more familiar to the GSU sample. Another possibility is the reduced cognitive load mentioned earlier. If 5- to 8-digit numbers did not induce sufficient cognitive burden, we are less likely to find evidence of restorative effect of nature or urban pictures. In that case, the effect we discovered in urban pictures could be something else, not the restorative effects on cognitive burden that we are initially searching for. A final possibility is that viewing pictures may not have the same effect as time in actual environments.

To explore the two possibilities, first, we ran the same experiment with 121 undergraduate students from the University of Alabama (AL) at the Interactive Decision Experiment (TIDE) Lab in the Culverhouse College of Business at AL. Summary statistics of the AL sample are presented in table 25. Table 26 shows the effect of long digits (cognitive load) on behavior before picture viewing. Most of the AL sample live on campus of the University of Alabama or close to campus, and they are surrounded by green and natural environment. If we find effects from nature but not urban with the AL sample, that may indicate existence of sample bias in GSU sample.

First, the rating of pictures of AL sample is lower than that of the GSU sample, for both nature and urban pictures, with nature still preferred to urban. The results in table 26 show that working with long-digits did not lower performance in attention-requiring task (Dots-mixed task); on the contrary, there were more correct responses in the Dots-mixed task under the long-digit scenario. Also, long-digit-induced burden did not lead to risk aversion, it resulted in less safe choices and more time later choices (more patience). These results raise the possibility again that our 5- to 8-digit numbers did not induce sufficient cognitive burden.

Second, mean comparison of responses before vs. after reviewing pictures across scenarios do not show any evidence of the effect from either nature or urban pictures (diff-in-dff) or from the comparison of the two picture sets (triple difference) in AL sample. Results from AL sample do not rule out the first possibility that subjects are more likely to be affected by the environment that they are more familiar with, and it raises the second possibility that cognitive burden induced was not sufficient. This will be incorporated in our next step to enhance the digit load or consider other methodologies to induce cognitive burden more effectively.

2.4 Conclusion and further steps

In this study, building upon previous studies in behavioral economics and environmental psychology, we examined the restorative effect of viewing nature or urban pictures on cognitive burden induced by memorizing numbers, and the consequent affect in choices involving attention-requiring tasks and risk and time preferences. Our results with GSU students show evidence that urban pictures help to reduce burden and decrease risk averse behavior with people who are familiar with urban environment.

However, the results reveal heterogeneity in individuals and groups and raise concerns on the level of cognitive burdens induced by designed long-digit numbers. We need to be aware about the possibility that the psychological effect from nature may not be universal. People may be more likely to relax in environment they are familiar with. On the other hand, the findings from current sample are not conclusive. To improve our understanding in the interactions of environment and human brains, further studies are required. In our next step, we will extend the length of numbers for inducing cognitive burden and reexamine our current conclusion with multiple samples.

Though surely no panacea for poverty and inequality, this research has the potential to extend our understanding in decision making involving risk and time preference under cognitive burdens. Insofar as the "environmental justice" literature has established that minorities and the poor are exposed to more pollution and less green space (Banzhaf 2012; Heynen et al. 2006), it also has the potential to elucidate one factor in economic inequality and the intergenerational transmission of disadvantage, in the US and globally. Although this study has not found a plausible solution, studies uncovering the myth of environment and human decision making has the potential to open the door to an intriguing pathway to more equitable as well as environmentally sustainable development.

Table 20 GSU ─ Summary Statistics before Picture Viewing

Percentage of Response	Pooled 1 digit	Pooled 5-8 Digit
Before picture viewing		
Digit Memorization	98.1%	69.5%
	5000	4840
Correct Dots	54.0%	51.3%
	2500	2372
Risk Safe Choice (gains)	49.1%	53.9%
	642	607
Risk Safe Choice (loss)	73.4%	71.2%
	661	605
Time later choice	47.6%	42.0%
	1197	1256

Note: number of observations below percentage of response of each task.

Table 21 GSU-Effect of Cognitive Load on Behavior - Pooled data of before picture viewing

Mean comparison (OLS) with robust standard error					
Dependent Variables:	Correct at Dots	Correct at Dots	Risk safe choice	Risk safe choice	Time later choice
	All sessions	Sessions 7-10	Gains	Losses	All sessions
long-digit number	-0.027**	-0.042**	0.048**	-0.021	-0.055**
t-value	[-1.91]	[-2.38]	[1.70]	[-0.85]	[-2.78]
Observations	4872	2785	1249	1266	2453

Notes: Dots-mixed task was enhanced from clicking the mouse to use keyboard at session 7. T-values are in Brackets. ** significant at 5%.

Table 22 GSU ─ Effect of Pictures - Pooled data of short- and long-digit numbers

Mean comparison (OLS) with robust standard error

Dependent Variables:	Correct at Dots	Risk safe choice	Risk safe choice	Time later choice
	All sessions	Gains	Losses	All sessions
After	0.038	-0.306	-0.312	-0.003
	[2.65]	[-1.06]	[-1.14]	[-0.12]
Nature	0.015	0.019	0.044	-0.041
	[1.07]	[0.67]	[1.73]	[-2.05]
After*Nature	-0.002	0.034	0.089**	-0.002
	[-0.08]	[0.83]	[2.48]	[-0.07]
Observations	9828	2463	2464	4904

Notes: T-values are in Brackets. ** significant at 5%.

Table 23 GSU ─ Summary Statistics

Percentage of Response	Nature Pictures: Short Digit	Nature Pictures: Long Digit	Urban Pictures: Short Digit	Urban Pictures: Long Digit
Before picture viewing				
Digit Memorization	97.9%	69.3%	98.3%	69.7%
	2520	2480	2480	2360
Correct Dots	56.0%	50.9%	52.1%	51.7%
	1230	1218	1270	1154
Risk Safe Choice (gains)	49.2%	55.4%	48.9%	52.1%
	327	319	315	288
Risk Safe Choice (loss)	73.5%	76.1%	73.7%	66.4%
	360	301	301	304
Time later choice	48.6%	37.2%	46.6%	47.1%
	603	642	594	614
After picture viewing				
Rate of Pictures	4.57		4	
Digit Memorization	98.5%	70.2%	98.3%	70.2%
	2520	2480	2480	2360
Correct Dots	59.6%	54.5%	55.5%	55.8%
	1267	1267	1225	1197
Risk Safe Choice (gains)	48.7%	56.3%	50.8%	43.7%
	308	321	297	288
Risk Safe Choice (loss)	77.8%	0.8%	73.2%	60.2%
	333	308	291	266
Time later choice	48.2%	36.1%	44.9%	48.5%
	606	579	662	604

Note: rate of pictures, 1=strongly dislike, 5=like a lot. Number of observations displayed below response percentage.

Table 24 GSU ─ Effect of Picture Viewing on Behavior

	Nature (diff-in-diff)			Urban (diff-in-diff)			Triple difference		
Dependent Variables:	Correct at Dots	Risk safe choice	Time later choice	Correct at Dots	Risk safe choice	Time later choice	Correct at Dots	Risk safe choice	Time later choice
After	0.036	0.021	-0.004	0.034	0.009	-0.018	0.035	0.028	-0.018
	[1.85]	[0.79]	[-0.14]	[1.74]	[0.31]	[-0.63]	[1.74]	[0.31]	[-0.63]
Long-digit	-0.051	0.038	-0.114	-0.004	-0.016	0.004	-0.004	-0.016	0.004
	[-2.54]	[1.41]	[-4.07]	[-0.19]	[-0.56]	[0.15]	[-0.19]	[-0.56]	[0.15]
After × long	-0.0002	0.017	-0.007	0.007	-0.087**	0.032	0.007	-0.087	0.032
	[-0.01]	[0.46]	[-0.18]	[0.24]	[-2.14]	[0.80]	[0.24]	[-2.14]	[0.80]
Nature							0.039	0.0067	0.019
							[1.95]	[0.25]	[0.68]
Nature× long							-0.047	0.053	-0.118
							[-1.65]	[1.38]	[-2.94]
After x nature							0.002	0.012	0.014
							[0.07]	[0.32]	[0.34]
Nature × long × after							-0.007	0.104**	-0.039
							[-0.18]	[1.89]	[0.056]
Constant	0.56	0.62	0.49	0.52	0.61	0.46	0.52	0.61	0.47
R-squared	0.004	0.003	0.01	0.002	0.006	0.0007	0.003	0.009	0.009
Obs.	4982	2577	2430	4846	2350	2474	9828	4927	4904

Notes: T-values are in brackets. ** significant at 5%. We also tried incorporating rate of picture in the analysis, nothing changed on treatment effect only that rate of picture is significant.

Table 25 Alabama ─ Summary Statistics

Percentage of Response	Nature Pictures		Urban Pictures	
	Short Digit	Long Digit	Short Digit	Long Digit
Before picture viewing				
Digit Memorization	97.3%	67.6%	96.9%	72.0%
	1240	1240	1200	1160
Correct Dots	67.1%	69.2%	63.6%	71.2%
	601	602	593	587
Risk Safe Choice (gains)	48.1%	47.7%	51.1%	35.0%
	162	132	172	163
Risk Safe Choice (loss)	69.0%	74.1%	71.5%	63.4%
	158	166	151	134
Time later choice	49.2%	61.2%	45.1%	52.2%
	319	340	284	276
After picture viewing				
Rate of Pictures	3.92		3.02	
Digit Memorization	98.6%	67.7%	98.2%	75.8%
	1240	1240	1200	1160
Correct Dots	71.7%	71.7%	69.8%	74.4%
	640	625	647	579
Risk Safe Choice (gains)	54.9%	51.7%	45.0%	36.1%
	173	149	131	133
Risk Safe Choice (loss)	67.1%	67.6%	75.9%	60.0%
	140	148	166	145
Time later choice	41.1%	56.3%	43.8%	52.8%
	287	318	256	303

Note: rate of pictures, 1=strongly dislike, 5=like a lot

Table 26 Alabama ─ Effect of Cognitive Load on Behavior

Mean comparison (OLS) with robust standard error

Dependent Variables:	Correct at Dots	Risk safe choice	Risk safe choice	Time later choice
		Gains	Losses	
long-digit number	0.049**	-0.098**	-0.009	0.099**
t-value	[2.56]	[-2.28]	[-0.24]	[3.47]
Observations	2383	629	609	1219

T-values are in brackets. ** significant at 5%.

3 CARBON TAXATION AND INDUCED INNVATION

3.1 Introduction

Providing the incentives to stimulate innovation is a key advantage of market-based carbon policies over other regulatory policies. Besides its static cost-effectiveness, which induces firms to equalize their marginal costs under current knowledge to the tax level, carbon taxes provide incentives to develop cleaner ways of production and novel ways to reduce emissions. Such innovation is crucial in the long run to reduce pollution-control costs (Jung et al. 1996). Moreover, technological innovations that help clean the environment also enable the company or industry to thrive and to be more competitive in global markets. The growth and better environment created by energy-related innovation will benefit people from all socioeconomic strata. That is probably why voters of all stripes are broadly supportive of technological innovation that can move us to a cleaner energy future (AMS, 2017).

However, we have very limited empirical evidence about the effect of market-based carbon policy on energy-efficient innovation. Previous studies either provide only theoretical results (Weber and Neuhoff, 2010), or focus on one just sector or industry, such as the auto industry (Aghion et al., 2016) or the German power sector (Rogge and Hoffmann, 2011). Since the seminal work of Popp (2002) on energy prices, we have learned that higher fuel prices induce firms to energy-efficient innovations. Insofar as a carbon tax will raise energy prices in general (bringing them up to their true social costs), I can examine the U.S. energy market and see if policy-induced increase in energy prices can stimulate energy-efficient innovations.

In this paper, I empirically examine the effect of a carbon tax on energy-efficient innovations with observed and simulated data. I intend to answer the following questions. First, given the current quality and stock of energy-related technology, what is the effect of higher energy prices (i.e. the tax) on innovation? Second, how would energy prices respond to a carbon tax implemented at a given point, in a partial equilibrium framework, while holding other market factors constant? And lastly, bridging the two parts, what is the effect of an energy tax on more energy-efficient innovation? The purpose of the paper is to provide empirical evidence on the ability of carbon taxes to stimulate or boost energy-efficient innovations and to provide insights into such effects at different tax levels. The findings can also inform general equilibrium studies on the overall effects of a carbon tax by providing indicators of the endogenous technological change.

I answer the proposed questions with the following steps. First, I replicate the work of Popp (2002) based on 1970-1994 data. In future work, I will extend this analysis to include more recent data. Following Popp, I specify innovation as a function of expected energy prices, the stock of knowledge, and other control factors. The expected energy prices depend on a weighted average of past prices. I also use estimated stocks of knowledge available to investors measure technological opportunities. As in Popp (2002), I use patent citations data to infer the quality of patents granted in term of usefulness.

Second, I simulate energy prices following a hypothetical carbon tax, imposed at a given year. I calculate the change in gross-of-tax fuel prices by adding the direct tax costs to observed prices. Specifically, suppose a carbon tax was implemented at a given year, I add proportionate cost of the tax to each type of fuel prices given its carbon dioxide emission coefficient from the U.S. Energy Information Administration (EIA). Then, the expected prices for the post-policy period is estimated based on the price updating model of Popp (2002), where the new expected prices depend on the post-tax price from the previous period and the price history.

Combining the two parts, I estimate the level of innovation that would have had occurred if a carbon tax was implemented in the past. My analysis permits one to explore the effects of various levels of carbon taxes. The results of these estimated effects of a carbon tax on energy-efficient innovation at different levels show the responsiveness of innovation to carbon taxes, especially in the long-run. For instance, given the average knowledge level and a carbon tax of $5 (per metric ton CO2) implemented in the year of 1971, there would be a threefold increase in the number of energy-efficient patents by the end of 1991. However, the incremental effects of raising the carbon tax levels are relatively small. The change in the number of energy-efficient patents under a carbon tax of $5 is similar to that with a carbon tax of $30.

3.2 Replication of Popp (2002)

Below I present the replication of Popp (2002) in two parts. In part one, I reproduce the results of productivity estimates for energy-efficient patents. In part two, I re-estimate the regression for the effects of energy prices on energy-efficient patents in the U.S.

3.2.1 Productivity of Patents and Stock of Knowledge

To study innovation, establishing the existing stock of knowledge on which inventors can build is essential. As Popp pointed out, if diminishing returns to research exit, increases in the current level of research and development (R&D) may make future R&D more difficult; on the other hand, technology accumulation is also important for innovative breakthrough. Hence, when we study the effect of energy prices on induced innovation, the stock of knowledge cannot be neglected. Also, since patents vary in their contribution and usefulness to future research, good estimates of the quality and productivity of patents become very important. Popp used patent citation data to construct productivity estimates. The following paragraph from Popp (2002) explains why citation data is a good source for this purpose:

> when a patent is granted, it contains citations to earlier patents that are related to the current invention. The citations are placed in the patent after consultations among the applicant, his or her patent attorney, and the patent examiner. It is the applicant's responsibility to list any related previous patents of which he or she is aware, and the examiner, who specialized in just a few patent classifications, will add other patents to the citations as well as subtracting any irrelevant patents cited by the inventor. Patent citations narrow the reach of the new patents by placing the patents cited outside the realm of the current patent, so it is important that all relevant patents be included in the citations. For the same reason, inventors have an incentive to make sure that no unnecessary patents are cited. As a result, the previous patents cited by a new patent should be a good indicator of previous knowledge that was utilized by the inventor (page 167).

The assumption Popp made was that the citations indicate a flow of knowledge. Also, he pointed out that *probability of citation* is a better indicator of the knowledge carried in a patent than a simple count of subsequent citations, because the raw number of citations to any patent depends on the total number of patents that follow. Below I replicate the productivity estimates using the patent citation data and model from Popp (2002). I begin with an introduction to the data; then I present the stages of estimation, and compares my results to Popp.

Table 27 presents a sample of the citation data used in Popp (2002). Table 28 lists the 12-energy technology group chosen by Popp due to data availability and quality (more details of data choice can be found in Popp, 2002). In table 27, patents granted from 1950 to 1990 were presented with the number of citations from 1974 to 1991. YEARAPP shows the citing year (the year patents cited earlier patents were applied). Column GROUP contains a character variable with the name of each energy technology group. PATAPPS shows the number of patents applications in each technology group in the years that citations were counted. Column GROUPNUM has a numeric index represents each energy technology group. PCTCTCLS records the probability of citation for patents within each group (details of calculation explained below). NGRNTCLS shows the number of patents in each technology group that were granted and YEARGRNT is the year those patents were granted. NCITES records the number of citations of the cited/citing cohort. As a result, the data were sorted by cited and citing year into groups of patents that could potentially cite each other. According to Popp, cross-groups citations were not considered because spillovers of knowledge do not necessarily impact R&D in the energy sector right away, including citations to all patents would complicate the induced innovation regression.

As mentioned above, PCTCTCLS records the probability of citation for patents within each group. Denote the granted year of cited patents as CTD, and the application year of citing patents as CTG, and denote each technology group as i, and the number of citations of each cohort in each technology group as $c_{i,CTD,CTG}$. Also, denote the number of potentially cited patents that were applied for in year CTD as $n_{i,CTG}$. Then the number of potentially citing patents granted in the year CTG as $n_{i,CTG}$, the probability of citation is written as:

(1) $p_{i,CTD,CTG} = \frac{c_{i,CTD,CTG}}{(c_{i,CTD})(n_{i,CTG})}$.

The probability of citation of each group cohort is modeled as in equation (2) by Popp:

(2) $p_{i,CTD,CTG} = \alpha_i \alpha_{i,CTD} \alpha_{CTG} exp[-\beta_1(CTG - CTD)]\{1 - exp[-\beta_2(CTG - CTD)]\} + \varepsilon_{i,CTD,CTG}$.

In equation (2), the first factor α_i stands for the effect of size of the technology group. It is a group fixed effect. The third factor $\alpha_{i,CTG}$ captures the effect of frequency with which patents applied for in the citing year cite earlier patents. It is a year fixed effects counting for any behavior change in citing over time. The second component, $\alpha_{i,CTD}$ presents the *productivity parameter*. Higher values of $\alpha_{i,CTD}$ imply that the patents are more likely to be cited and probably that the knowledge embodied in those patents is particularly useful. The two exponential distributions model the flow of knowledge. β_1 represents the rate of decay of knowledge as it become obsolete, and β_2 represents the rate at which newly produced knowledge (newly patented innovation) diffuses through society.

Using the patent citation data from Popp (2002), I re-estimate the productivity of patents in each technology group over years in equation (2) with nonlinear least squares. My results are presented in Figure1. As in Popp (2002), year 1970 of each technology group was normalized to 1. The results in Figure 2 are comparable to Figure 3 in Popp (2002).

In general, my productivity estimates are very close to those in Popp (2002). The productivity pattern over time is very consistent with the findings in Popp. However, there are a few estimates slightly different from Popp. First, my productivity estimates of solar energy, solar batteries, heat exchange, and continuous casting around 1976 are lower than those in Popp. Also, my estimates of waste heat after 1988 is slightly lower than that in Popp. These differences may be due to the ways that different statistic programs handle nonlinear least squares.[6]

Next, knowledge stocks of each energy group are constructed. As in Popp, a stock of patents weighted by the productivity estimates are measured as in equation (3).

(3) $K_{i,t} = \sum_{s=0}^{t} \alpha_{i,s}\, PAT_{i,s} exp[-\beta_1(t-s)]\{1-exp[-\beta_2(t-s)]\}$.

As specified in equation (3), the accumulated stock of knowledge at a certain time is measured by the sum of patents granted in previous years $PAT_{i,s}$, weighted by the productivity factor $\alpha_{i,s}$, and the time decay factors $exp[-\beta_1(t-s)]$, and diffusion factors $1-exp[-\beta_2(t-s)]$. And the unweighted stocks are calculated in the same way without the productivity weights $\alpha_{i,s}$. I present my results of weighted stock of knowledge of each energy technology group over years in Figure 3. These results are comparable to Figure 4 in Popp (2002). As in Popp, the stocks of year 1970 were normalized to 1. In Figure 3 the dashed lines are unweighted stocks and solid lines represent the weighted stocks.

[6] I programmed the model in MATLAB and Popp's coding was in GAUSS.

Compared to Figure 4 in Popp (2002), my weighted stocks of knowledge are very consistent with those in Popp. However, there are a few mismatches in the unweighted stocks. This discrepancy might be due to inclusion/exclusion of certain years.[7] The weighted stocks contain the information of quality and usefulness of patents, but the unweighted stocks do not. Only the weighted stocks of knowledge are used in the second stage below when examining the effects of energy prices on patents.

3.2.2 Energy Prices and Innovation

After estimating of stock of knowledge, I turn to the analysis of patents and energy prices. The model from Popp was specified as follows:

$$(4)\ log(\frac{EPAT_{i,t}}{TOTPAT_t}) = \varphi_i + \gamma(1-\lambda)logP^*_{E,t} + \theta logK_{i,t-1} + \eta(1-\lambda)log\boldsymbol{Z}^*_{i,t} + \lambda^t\mu^0 + \varepsilon_{it},$$

where $P^*_{E,t} = P_{E,t} + \lambda P_{E,t-1} + \lambda^2 P_{E,t-2} + \ldots + \lambda^{t-1}P_{E,1}$, and

$\boldsymbol{Z}^*_{i,t} = Z_{i,t} + \lambda Z_{i,t-1} + \lambda^2 Z_{i,t-2} + \ldots + \lambda^{t-1}Z_{i,1}$, with $i = 1,\ldots,11;\ t = 1,\ldots,20.$

In this model, the dependent variable is $EPAT_{i,t}$, the number of successful nongovernment U.S. patent applications for technology *i* in year *t*, divided by $TOTPAT_t$, the total number of successful nongovernment U.S. patent applications in the same year. $P^*_{E,t}$ is the expected energy price in year *t*, and it depends on the weighed average of past prices. $\boldsymbol{Z}^*_{i,t}$ is a vector that contains the expected R&D spending by the U.S. Department of Energy and some other group-specific explanatory variables as listed in Popp (2002) (see page 164). As with energy prices, the expected values are included in the regression.

$K_{i,t-1}$ represents the previous stock of knowledge of group *i* as estimated in section 2.1. μ^0 stands for the truncation remainder, because an infinite series of past independent variables is not

[7] I used all available years (1899-1995) when calculating the unweighted stocks of knowledge. It is unclear which years Popp adopted in the paper.

possible. By specification, $\gamma(1-\lambda)$ represents the short-run price elasticity of energy innovation, and γ is the long-run price elasticity. Lagged party of the president and lagged government R&D are used as instruments for government R&D. A time trend and lagged values of other exogenous variables are used as instruments for the knowledge stock[8].

Using the U.S. patent data from 1971 to 1991 as in Popp (2002), and with the stock estimates from section 2.1, I re-estimate the model in equation (4) with Generalized Method of Moments (GMM). First, I find the estimates of γ given the first moment defined by equation (4), then substitute the estimated value of γ into the model and run the linear model to obtain statistical t values and confidence intervals for the coefficients.

My results for equation (4) are presented in table 29. The results from Popp (2002) are copied from his table 4 and shown in table 29 for comparison. The magnitude of the effect of energy prices is very close to that obtained by Popp and is significant. However, I obtained a higher impact of the weighted stock of patents, which is highly significant. To reiterate Popp's points on these results: "the price elasticities found suggest the reaction of the research community to a change in policy, such as a carbon tax, will be swift, and that higher prices would quickly lead to a shift toward environmentally friendly innovation. In addition, the positive knowledge stock coefficients suggest that the usefulness of the existing base of knowledge is important to inventors – inventors 'stand on the shoulders' of their predecessors."

In this paper, I focus on the short- and long-run effect of energy prices. Given the estimates of short-run price elasticity of energy innovation at 0.05 and the long-run price elasticity as 0.569, I calculate the increased innovations generated by price change in table 30. From year t to $t+1$,

---

[8] Popp collected 25 variables as potential instruments for knowledge stocks, including macroeconomic indicators and technological group specific outputs. It is unclear which of those were applied. In this paper, I used all possible variables from Popp's list. Detailed list is available upon request.

suppose price increase by the percentage specified in table 30, correspondingly, the percentage change in the ratio of energy-efficient patents (average of the 11 technology groups) to total patents in the U.S. was calculated. For example, suppose the total patents at year t is 1000, and there are 50 patents in the energy-efficient sector, the number of patents will increase to 50.76 in the short-run and 59.34 in the long-run, when energy prices are expected double.

3.3 Carbon Tax and Energy Prices

In this section, I construct post-CO2 tax energy prices. Energy prices in equation (4) were obtained from U.S. Energy Information Administration (EIA) as in Popp (2002). The EIA constructs average energy prices for major production and consumption sectors by dividing total expenditure by total consumption.[9] The total average price per year constructed by EIA was used in the regression.

Examining the average energy price levels after a national carbon tax normally requires a general equilibrium analysis incorporating the dynamics of demand and supply, in which technological innovation is endogenous. Since this paper focuses on the partial equilibrium effect of energy prices on innovation, I adopt a less complex and straightforward method by transforming the effective price of tax to different type of energy by their CO2 emission coefficients. This method is the same as that presented by Hafstead & Picciano (2017).

I first obtain the CO2 emission coefficient from EIA of different types of fuels as in table 31. The emission coefficients measure the weight of CO2 released per million British thermal unit of fuel. The costs added to each type of fuel can be calculated as coefficients/1000*tax at each tax level specified. The hypothetical tax levels are in current dollars, and they are converted by CPI

[9] See details at https://www.eia.gov/state/seds/data.php?incfile=/state/seds/sep_prices/total/pr_tot_US.html&sid=US

when applied to previous years. The tax levels are chosen based on previous carbon tax that has been considered or implemented worldwide.[10]

Energy types in table 31 are matched with types of the EIA coefficients and prices. The categories of the coefficients and prices are fairly consistent, with a few exceptions. For Hydrocarbon gas liquids (HGL), I take the average coefficients of propane, butane and flared natural gas. And for wood and waste, I take the average coefficients of municipal solid waste and waste oil. Then, the added costs from each type of fuel are weighted by the ratio of fuel consumption to total consumption. And the aggregate weighted costs are added to the total energy price to obtain the post-tax energy prices. Since all energy prices are in 1987 constant dollars, all added costs are converted to 1987 constant dollars with the CPI index. Table 32 presents total energy price adopted by Popp and the aggregate weighted added costs by different tax levels. All prices in table 32 are in 1987 constant dollars except the tax levels in the header.

While calculating the consumption weights for added costs, electricity consumption was not included. In the U.S., electricity is generated by other primary fuels (such as coal, natural gas and nuclear energy, and other renewable resources such as hydropower, biomass, wind, geothermal and solar power). Since the emission coefficients of related primary fuels are incorporated, it would be double counting to add the costs of electricity. Also, Coke is counted for imported only (not exported) for simplicity.

[10] In 2012, Australia proposed a carbon tax of AUD $23 per tonne of emitted CO2. New Zealand proposed NZ $15 per tonne of CO2 in 2005. The European Commission suggested €4 to €30 per tonne of CO2 to EU countries. And British Columbia implemented a tax of $10 per tonne of CO2 emission in 2008.

3.4 Carbon Tax and Induced Innovation

Combining the estimates from section 2 and 3, now I present the results of innovation that could be generated by carbon taxes. First, I use the formula of expected prices in equation (4) to obtain the expected price when a carbon tax is introduced. For instance, as in table 32, the tax costs associated with various tax levels (total added costs) are added to the total energy prices adopted by Popp to obtain the post-tax energy prices. Then, expectations of future prices are formed based on the post-tax energy prices following the adaptive expectation model as specified in equation (4). The assumption is that individuals are likely to take the post-tax prices in the past and put weights on them to predict price changes in the future. For instance, if a carbon tax was implemented in 1970, how would people in 1980 form expectations of energy prices for 10 years later, as of 1990? Given the added costs from the carbon tax in table 32, energy prices observed from 1970 to 1980 would have increased by the total added costs. Adding the total added costs of various tax levels to the total energy prices used in Popp, I obtain the post-tax energy prices. And these post-tax energy prices from 1970-1980 would be the observations to people in 1980. According to adaptive expectation theory, these would be the information people took into consideration when they speculated future price changes.[11] That is, the model assumes people treat the tax as a repeated shock, the importance of which they discount especially in the short-run as their expectations are that prices will revert to their recent mean. An alternative would be to first predict expectations of the net-of-tax price using the adaptive expectations model, and then add

[11] Please note that this method assumes away the dynamics in energy price changes. As people/firms expect higher prices and change their consumption/production behavior, the change in supply and demand and the technological innovations will all affect future prices. Hence, the price realized may differ from the estimated post-tax prices based in historical data.

the tax to their expectation. That approach would assume people treat the tax component of the price is a permanent shock.

Second, I apply the energy innovation price elasticity to obtain the short-run and long-run changes in innovation that are triggered by the price change, holding everything else constant. The partial effects of additional increase in energy prices from the tax are directly obtained by multiplying the price elasticity with any price changes. Results are shown in table 33 and 34 for the short-run price elasticity and long-run price elasticity, respectively. I arbitrarily chose the years of 1981-1991 to exemplify the long-run effects. In tables 33 and 34, the second column presents the number of patents without tax of the technology group "using waste as fuel" as an example of the counterfactual group. The effects of taxes in terms of the number of patents generated can be obtained by subtracting the number of patents at tax zero from the number of patents post tax at a specific level. Please note the increase in number of patents post tax in table 33 and 34 is for one technology group only, the gross effect for all energy-efficient patents is much larger. For instance, the total effect for the 11 technology groups in study would be 11 times larger based on the average price elasticity estimated.

The changes in the number of new patents following price changes illustrate the effectiveness and significance of the price elasticity of energy innovations. Firstly, the short-run effect on one single technology group seems small. This is a result from the small magnitude of the tax increasement relative to the energy price levels, especially after multiplying with factors in emission coefficients. Secondly, in the long-run, the effect on one technology group becomes more obvious. This is consistent with the time that innovations would demand. It is important to remember that there are more than 11 energy efficient technology groups and the estimated price elasticity only counted the average effect on the chosen 11 groups due to data availability. Hence, the actual gross effect is likely to be much larger than my calculation and it might occur faster.

3.5 Conclusion and Future Steps

In the analysis so far, I have replicated the results of Popp (2002) with 1970-1994 data and used the estimated coefficients on energy prices to quantify the number of new energy-efficient innovation that would be stimulated by various levels of carbon taxes. The results in this paper only consider the partial-equilibrium causal effects of carbon taxes and do not analyze the dynamic changes in demand and supply when examining the price increase post a carbon tax. Still, the results are informative for policy makers and researchers working on the dynamic effects of carbon taxes where induced innovations play a key role in determining the future costs of emission abatements and energy consumption. And it provides a useful starting point for future analysis under a general-equilibrium framework.

The next step is to update the estimates with more recent patent data and to obtain innovation changes for recent and future periods. First, I will update the estimates of technology stock and price coefficients with up-to-data patents and prices data. Second, I will conduct out-of-sample simulation of expected energy prices for future years with an ARMA (Auto-regressive moving-average) time series model based on historical prices. The ARMA model regresses prices on its own lagged values and adjusts the average level of prices overtime to minimize the remaining noise. It assumes that future energy prices would largely depend on past prices. The model and its variation have been widely used in predicting electricity and crude oil prices (Cuaresma et al., 2004, Liu and Shi, 2013). Lastly, I will quantify the innovation that would emerge if a carbon tax is imposed now and in the future. The results can be used as a reference points for predicting potential innovation benefits of carbon tax policies.

Table 27 Patents Citation Data (Popp 2002)

Obs	YEAR APP	GROUP	PAT APPS	GROU PNUM	PCTCTCLS	YEARGR NT	NGRNTC LS	NCIT ES
1	1974	coalliq	63	1	0	1950	2	0
2	1975	coalliq	58	1	0	1950	2	0
3	1976	coalliq	115	1	0	1950	2	0
4	1977	coalliq	106	1	0	1950	2	0
…	…	…	…	…	…	…	…	…
580	1989	coalliq	29	1	0.007	1988	26	5
581	1990	coalliq	23	1	0.012	1988	26	7
582	1991	coalliq	17	1	0.000	1988	26	0
583	1990	coalliq	23	1	0.011	1989	38	10
584	1991	coalliq	17	1	0.008	1989	38	5
585	1991	coalliq	17	1	0.009	1990	20	3
586	1974	coalgas	51	2	0	1950	2	0
587	1975	coalgas	49	2	0	1950	2	0
588	1976	coalgas	65	2	0	1950	2	0
589	1977	coalgas	63	2	0	1950	2	0
590	1978	coalgas	70	2	0	1950	2	0
591	1979	coalgas	53	2	0	1950	2	0
…	…	…	…	…	…	…	…	…
6435	1991	contcast	107	11	0	1990	164	11

Table 28 Energy Groups in Popp (2002)

Supply Technology	
coalliq	Coal liquefaction: producing liquid fuels
coalgas	Coal gasification: producing gaseous fuels
solareng	Solar energy
solrbtry	Batteries for storing solar energy
fuelcell	Fuel cells
wstfuel	Using waste as fuel
Demand Technologies	
wstheat	Recovery of waste heat for energy
heatx165	Heat Exchange, general
heatpump	Heat Pumps
stireng	Stirling engines
contcast	Continuous casting processing of metal

Table 29 Induced Innovation Regression Results

Independent Variable	Weighted Stock of Patents	Popp (2002)
Constant	-2.120	-7.311
	[-37.783]	[-46.625]
Energy Prices: $\gamma(1-\lambda)$	0.050	0.060
	[3.653]	[2.852]
Lagged knowledge stock: θ	2.044	0.838
	[31.16]	[72.3231]
Government R&D: $\eta(1-\lambda)$	-0.002	-0.009
	[-0.403]	[-1.741]
Truncation error: μ^0	-1.027	-1.203
	[-7.795]	[-5.054]
λ	0.912	0.829
long-run energy elasticity: γ	0.569	0.354
Long-run government R&D elasticity: η	-0.023	-0.025
Number of technology groups	11	11

Note: T values are in parentheses.

Table 30 Price Increase and Innovation

Percentage increase in expected price	5%	10%	15%	20%	50%	100%
Short-run						
Percentage change in patents ratio	0.11%	0.21%	0.30%	0.40%	0.88%	1.52%
At total PAT=1000, ratio=5%						
# of PAT post price increase	50.05	50.10	50.15	50.20	50.44	50.76
Long-run						
Percentage change in patents ratio	1.21%	2.38%	3.51%	4.61%	10.54%	18.68%
At total PAT=1000, ratio=5%						
# of PAT post price increase	50.61	51.19	51.76	52.30	55.27	59.34

Table 31 Fuel Emissions and Tax Costs

Energy Type	CO2 coefficient	Cost added w/tax=$5	tax=$10	tax=$15	tax=$20	tax=$25	tax=$30
Coal (all type)	95.35	$0.48	$0.95	$1.43	$1.91	$2.38	$2.86
Coke	114.12	$0.57	$1.14	$1.71	$2.28	$2.85	$3.42
Natural Gas	53.07	$0.27	$0.53	$0.80	$1.06	$1.33	$1.59
Home Heating and Diesel Fuel (Distillate)	73.16	$0.37	$0.73	$1.10	$1.46	$1.83	$2.19
HGL	50.00	$0.25	$0.50	$0.75	$1.00	$1.25	$1.50
Jet Fuel	70.90	$0.35	$0.71	$1.06	$1.42	$1.77	$2.13
Motor Gasoline	71.30	$0.36	$0.71	$1.07	$1.43	$1.78	$2.14
Residual Fuel Oil	78.79	$0.39	$0.79	$1.18	$1.58	$1.97	$2.36
Petroleum other	72.62	$0.36	$0.73	$1.09	$1.45	$1.82	$2.18
Nuclear Fuel	0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
Wood and Waste	68.47	$0.34	$0.68	$1.03	$1.37	$1.71	$2.05

Note: CO2 coefficient is measured in kilogram CO2 per million btu of fuel. Tax price is per metric ton CO2 in current (2018) dollars. CO2 coefficients are estimates from U.S. Energy Information Administration. Coefficients vary slightly across time but the difference since 1970 is minor.

Table 32 Energy Groups in Popp (2002)

year	Total Energy Price (Popp)	TTL. Added Cost w/tax=$5	tax=$10	tax=$15	tax=$20	tax=$25	tax=$30
1970	4.56	0.16	0.32	0.48	0.64	0.79	0.96
1971	4.62	0.16	0.32	0.48	0.63	0.78	0.95
1972	4.62	0.16	0.32	0.47	0.63	0.78	0.95
1973	4.81	0.16	0.32	0.48	0.63	0.79	0.95
1974	6.21	0.16	0.31	0.47	0.63	0.78	0.95
1975	6.57	0.16	0.31	0.47	0.63	0.78	0.94
1976	6.68	0.16	0.31	0.47	0.63	0.78	0.94
1977	6.97	0.16	0.31	0.47	0.63	0.78	0.94
1978	6.94	0.16	0.31	0.47	0.62	0.77	0.94
1979	7.81	0.16	0.31	0.47	0.63	0.78	0.94
1980	9.39	0.16	0.31	0.47	0.63	0.78	0.94
1981	10.09	0.16	0.31	0.47	0.63	0.78	0.94
1982	10.05	0.16	0.31	0.47	0.62	0.77	0.94
1983	9.67	0.16	0.31	0.47	0.63	0.78	0.94
1984	9.28	0.16	0.31	0.47	0.63	0.78	0.94
1985	8.97	0.16	0.31	0.47	0.62	0.78	0.93
1986	7.62	0.16	0.31	0.47	0.62	0.77	0.93
1987	7.37	0.16	0.31	0.47	0.62	0.77	0.93
1988	7.03	0.15	0.31	0.46	0.62	0.77	0.92
1989	7.08	0.15	0.31	0.46	0.61	0.76	0.92
1990	7.37	0.15	0.30	0.46	0.61	0.76	0.91
1991	7.06	0.15	0.30	0.45	0.60	0.75	0.91

Note: Total energy price (Popp) are observed historical energy price without carbon tax in the U.S. market. Level of taxes are in current dollars. Other prices and costs are in constant 1987 dollars.

Table 33 Carbon Tax and Innovation (short-run price elasticity)

Carbon tax per metric ton CO2	$0	$5	$10	$15	$20	$25	$30
Year	# of PAT (waste fuel)	# of PAT post tax	#	#	#	#	#
End of 1971	53.0	53.1	53.2	53.2	53.3	53.4	53.5
1972	52.0	52.1	52.2	52.2	52.3	52.4	52.5
1973	52.0	52.1	52.2	52.2	52.3	52.4	52.5
1974	49.0	49.1	49.2	49.2	49.3	49.4	49.4
1975	29.0	29.1	29.1	29.2	29.3	29.3	29.4
1976	32.0	32.1	32.1	32.2	32.3	32.3	32.4
1977	34.0	34.1	34.1	34.2	34.3	34.3	34.4
1978	41.0	41.1	41.1	41.2	41.2	41.3	41.4
1979	40.0	40.1	40.1	40.2	40.2	40.3	40.3
1980	50.0	50.1	50.1	50.2	50.2	50.3	50.3
1981	44.0	44.1	44.1	44.2	44.2	44.3	44.3
1982	58.0	58.1	58.1	58.1	58.2	58.2	58.3
1983	50.0	50.1	50.1	50.1	50.2	50.2	50.3
1984	44.0	44.0	44.1	44.1	44.2	44.2	44.3
1985	46.0	46.1	46.1	46.2	46.2	46.2	46.3
1986	61.0	61.1	61.1	61.2	61.2	61.2	61.3
1987	83.1	83.1	83.2	83.2	83.3	83.3	83.3
1988	69.1	69.2	69.2	69.3	69.3	69.4	69.4
1989	84.5	84.5	84.6	84.6	84.7	84.7	84.8
1990	102.4	102.5	102.5	102.6	102.6	102.6	102.7
1991	97.5	97.6	97.6	97.7	97.7	97.8	97.8

Note: The column of tax=$0 contains number of patents for technology group waste fuel.

Table 34 Carbon Tax and Innovation (long-run price elasticity)

Carbon tax per metric ton CO2	$0	$5	$10	$15	$20	$25	$30
Year	# of PAT	# of PAT post tax	#	#	#	#	#
1987	83.1	83.6	84.2	84.7	85.3	85.8	86.3
1988	69.1	69.7	70.2	70.8	71.3	71.9	72.4
1989	84.5	85.1	85.6	86.2	86.7	87.2	87.8
1990	102.4	103.0	103.5	104.1	104.6	105.2	105.7
1991	97.5	98.1	98.7	99.2	99.8	100.3	100.9

Note: The column of tax=$0 contains number of patents for technology group waste fuel.

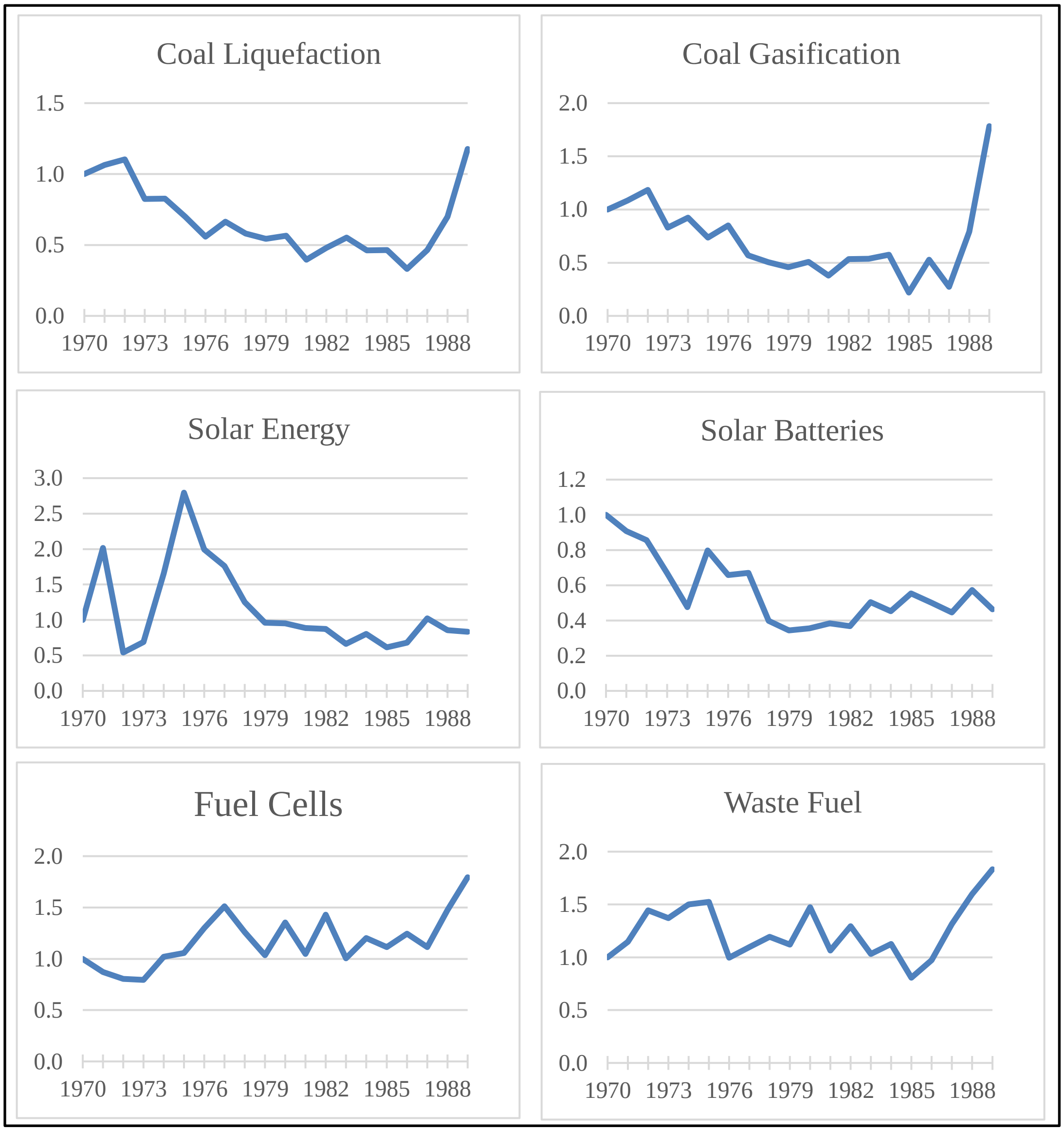

Figure 2 Estimated Productivity of Patents

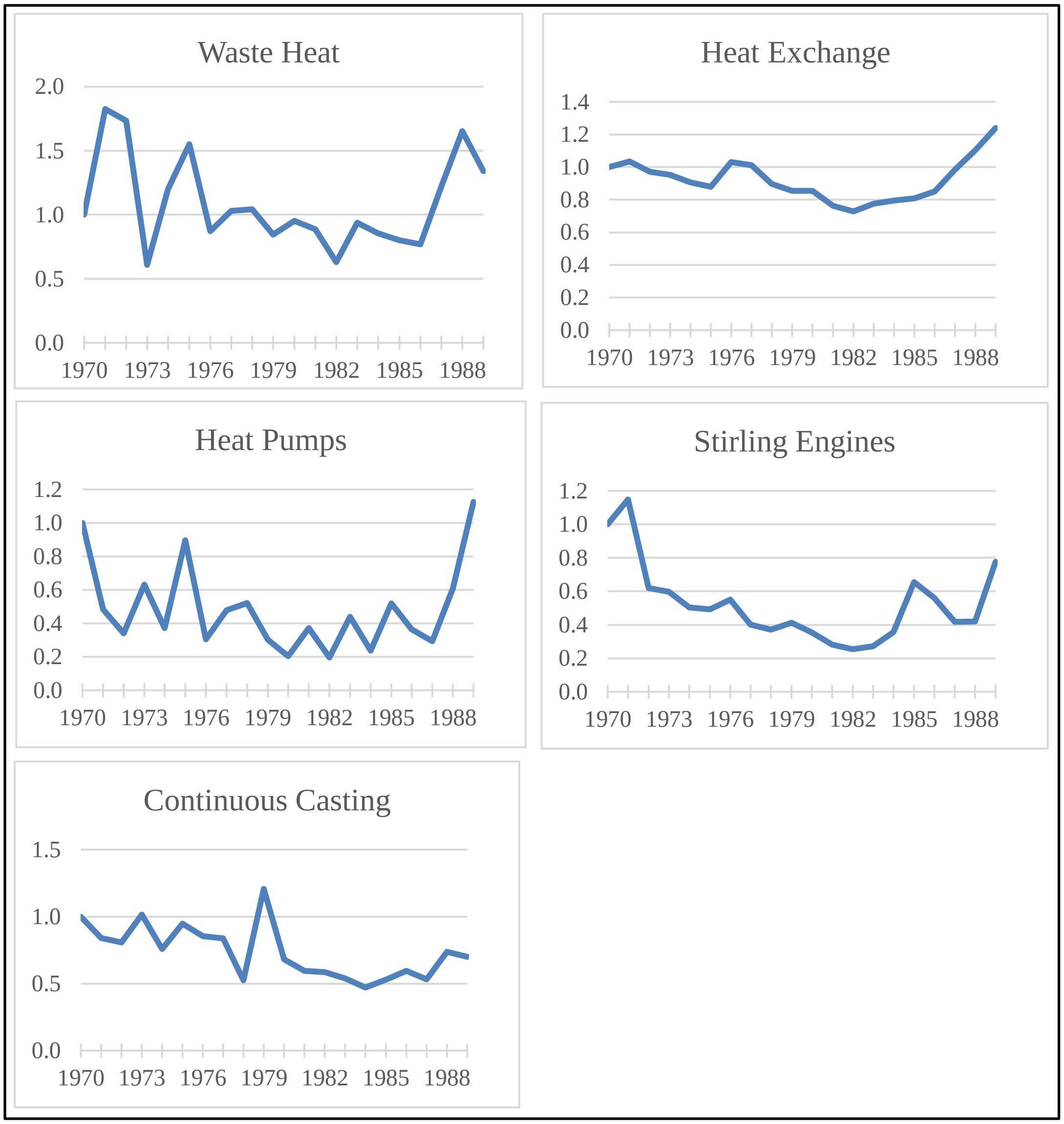

Figure 3 Estimated Productivity of Patents-continued

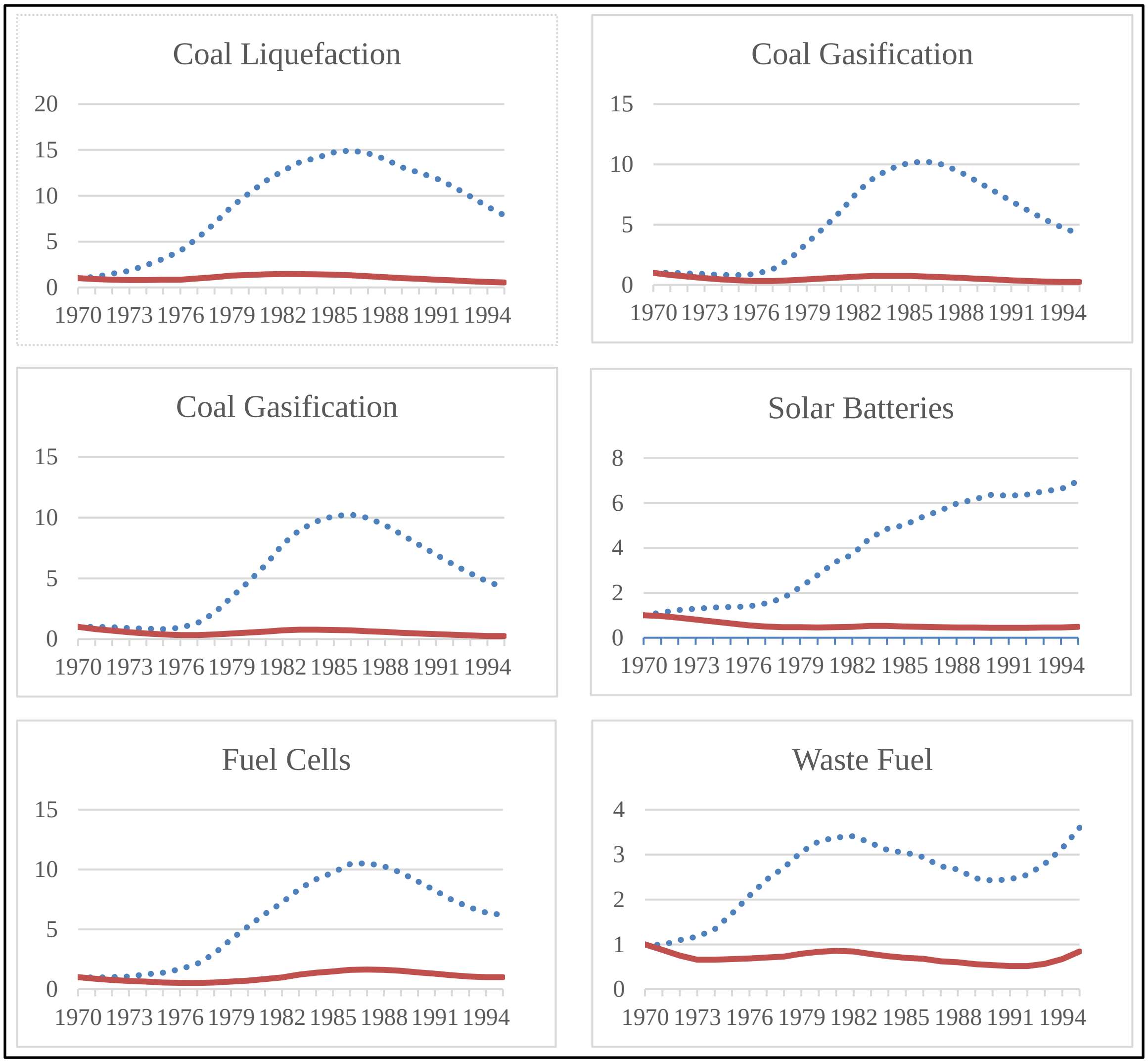

Figure 4 Stock of Knowledge

Figure 5 Stock of Knowledge-continued

REFERENCES

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

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R. and Van Reenen, J., 2016. Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, *124*(1), pp.1-51.

Alliance for Market Solutions (2017). *Carbon Tax Policy: A Conservative Dialogue on Pro-Growth Opportunities.* Alex Brill, Editor.

Andreoni, J. and Sprenger, C., 2012. Estimating time preferences from convex budgets. *The American Economic Review*, *102*(7), pp.3333-3356.

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. 2012. *The Political Economy of Environmental Justice*. Stanford Univ. Press.

Berman, Marc G., John Jonides, and Stephen Kaplan. 2008. The Cognitive Benefits of Interacting with Nature. *Psychological Science* 19(12): 1207-12.

Berto, Rita. 2005. Exposure to restorative environments helps restore attentional capacity. *Journal of Environmental Psychology* 25(3): 249-59.

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

Bowler, Diana E., Lisette M. Buyung-Ali, Terry M. Knight, and Andrew S. Pullin. 2010. "A Systematic Review of Evidence for the Added Benefits to Health of Exposure to Natural Environments."*BMC Public Health* 10(1): p.456.

Bratman, Gregory N., J. Paul Hamilton, and Gretchen C. Daily. 2012. "The Impacts of Nature Experience on Human Cognitive Function and Mental Health." *Annals of the New York Academy of Science* 1249: 118-36.

Bratman, Gregory N., J. Paul Hamilton, Kevin S. Hahn, Gretchen C. Daily, and James J. Gross. 2015. "Nature Experience Reduces Rumination and Subgenual Prefrontal Cortex Activation." *PNAS* 112(28): 8567-72.

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.

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.

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.

Cuaresma, J.C., Hlouskova, J., Kossmeier, S. and Obersteiner, M., 2004. Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy*, *77*(1), pp.87-106.

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

Deck, Cary, and Salar Jahedi. 2015. "The Effect of Cognitive Load on Economic Decision Making:A Survey and New Experiments." *European Economic Review* 78: 97-119.

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.

Hafstead, Marc and Picciano, Paul 2017. Great Ideas in Environmental Economics, published by Resource for the Future: https://www.resourcesmag.org/common-resources/calculating-various-fuel-prices-under-a-carbon-tax/

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.

Hartig, Terry, Gary W. Evans, Larry D. Jamnerc, Deborah S. Davis, and Tommy Garling. 2003. Tracking restoration in natural and urban field settings. *Journal of Environmental Psychology* 23:109-23.

Hartig, Terry, Richard Mitchell, Sjerp de Vries, and Howard Frumkin. 2014. "Nature and Health." *Annual Review of Public Health* 35: 207-28.

Hartig, Terry, and Henk Staats. 2006. "The Need for Psychological Restoration as a Determinant of Environmental Preferences." *Journal of Environmental Psychology* 26: 215-26.

Hartig, Terry, Lars Nyberg, Lars-Göran Nilsson, and Tommy Gärling. 1999. "Testing for Mood CongruentRecall with Environmentally Induced Mood." *Journal of Environmental Psychology* 19(4): 353-67.

Haushofer, Johannes, and Ernst Fehr. 2014. "On the Psychology of Poverty." *Science* 344(6186): 862-7.

Heynen, Nik, Harold A. Perkins, and Parama Roy. 2006. "The Political Ecology of Uneven Urban Green Space." *Urban Affairs Review* 42: 3-25.

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

Jung, Chulho, Kerry Krutilla, and Roy Boyd. "Incentives for advanced pollution abatement technology at the industry level: An evaluation of policy alternatives." *Journal of environmental economics and management* 30, no. 1 (1996): 95-111.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus, and Giroux.

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.

Liu, H. and Shi, J., 2013. Applying ARMA–GARCH approaches to forecasting short-term electricity prices. *Energy Economics*, *37*, pp.152-166.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 341(6149): 976-80.

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.

Mullainathan, Sendil, and Eldar Shafir. 2013. *Scarcity: The New Science of Having Less and How it Defines our Lives.* New York: Picador.

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.

Popp, D., 2002. Induced innovation and energy prices. *American economic review*, *92*(1), pp.160-180.

Rogge, K.S., Schneider, M. and Hoffmann, V.H., 2011. The innovation impact of the EU Emission Trading System—Findings of company case studies in the German power sector. *Ecological Economics*, *70*(3), pp.513-523.

Schilbach, Frank, Heather Schofield, and Sendhil Mullainathan. 2016. "The Psychological Lives of the Poor." *American Economic Review* 106(5): 435-40.

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

Shah, Anuj K., Sendhil Mullainathan, and Eldar Shafir. 2012. "Some Consequences of Having too Little." *Science* 338(6107): 682-5.

Standal, Dag, and Frank Asche. 2018. "Hesitant Reforms: The Norwegian Approach towards ITQ's." *Marine Policy*, *88*, pp.58-63.

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.

Thompson, Catharine Ward, Jenny Roe, Peter Aspinall, Richard Mitchell, Angela Clow, and David Miller. 2012. "More Green Space is Linked to Less Stress in Deprived Communities: Evidence from Salivary Cortisol Patterns." *Landscape and Urban Planning* 105(3): 221-9.

Ulrich, R.S., R.F. Simons, B.D. Losito, E. Fiorito, M.A. Miles, and M. Zelson. 1991. "Stress Recovery during Exposure to Natural and Urban Environments." *Journal of Environmental Psychology* 11(3): 201-30.

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.

Wells, Nancy M. 2000. "At Home with Nature: Effects of 'Greenness' on Children's Cognitive Functioning." *Environment and Behavior* 32(6): 775-5.

White, Mathew P., Ian Alcock, Benedict W. Wheeler, and Michael H. Depledge. 2013. "Would You Be Happier Living in a Greener Urban Area? A fixed-Effects Analysis of Panel Data." *Psychological Science*: 920-8.

APPENDIX

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

<u>*Proof*</u>

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. However, given the nonlinearity in production function, we do not guarantee a unique TOC equilibrium as in Carvajal et al. (2013), that means when data passes the linear program, it is not guaranteed that data was generated under a TOC regime.