Trading for Economies of Scale: The Cost Efficiency of Cap-and-Trade in Common-Pool Resources

Phuong Ho[*]

June 27, 2022

Abstract

Cap-and-trade (CAT) has been considered cost-effective because trading provides a means for the cap to flow towards their highest-valued use. I offer a new perspective of CAT: Cost efficiency owing to economies of scale. That is, if the average cost of a firm is decreasing, trading will allow the firm to expand their operation, thereby producing at lower per-unit costs. To examine this hypothesis, I overcome the unobserved-cost challenge by introducing a method to estimate output elasticity of costs and infer economies of scale using data on output and input quantity. I combine this method with various approaches of estimating a production function and the identification strategy of differences-in-differences framework that exploits the policy transition from non-tradable cap to cap-and-trade in Norwegian cod fishery. Results show vessels had economies of scale before trading and vessels acquiring caps expanded their operation and moved toward the minimum average cost levels. When decomposing the output growth into factors of economies of scale that capture the cost efficiency and the component due to productivity change that captures the cost effectiveness, I find economies of scale played the main role in the first few years after a vessel acquires caps, whereas productivity improvement dominates the output growth afterwards. The finding of gains owing to economies of scale suggests the gains of CAT and resource rent may be underestimated if only the cost effectiveness due to productivity heterogeneity is evaluated.

Keywords: cap-and-trade, economies of scale, productivity, production function

JEL Classification: D22, D24, L11, Q22, Q28

---

[*] Cornell University and SNF—Centre for Applied Research at Norwegian School of Economics. Email: phuongho@cornell.edu. I thank Ashley Langer, Derek Lemoine, Alex Shcherbakov, Mauricio Varela, and seminar participants at SNF, University of Arizona, Maastricht University, Dyson School at Cornell University, AERE@SEA 2021, IIOC 2022, and CREEA 2022 for thoughtful discussions and comments. I thank Norwegian Directorate of Fisheries for providing access to the data as well as Tove Aasheim, Anita Kjeilen, Linda Nøstbakken, and Mads Wold for many clarifications on the dataset structure and Norwegian fishery regulations.

1 Introduction

Cap-and-trade (CAT) has increasingly become a popular choice by policy makers to reduce air emissions. Compared to command-and-control (CAC) approaches, cap-and-trade provides a means for the cap to flow towards their highest-valued use. Hence, trading will equalize the marginal abatement costs across firms in the industry to help achieve the second-best social welfare solution, as known as the equimarginal principle or the cost effectiveness of CAT. With similar reasons, CAT has been also introduced in common-pool resource management to avoid over-exploitation and maintain conservation; see Tietenberg (2003); Costello et al. (2010). For example, in fishery, a system known as individual transferable quotas (ITQs) gives a vessel the right to catch up to a ceiling amount (quota) that is tradable in the market. However, many territories, although recognizing the importance of caps in protecting the common resource, oppose to the tradable part of the system. They concern trading will lead to consolidation of quotas and outputs to certain big firms that will abuse market power. This paper studies the effects of CAT, as compared to the non-tradable cap system, on individual performance in the fishing market to test the market power hypothesis.

Furthermore, I offer a new perspective of CAT: cost efficiency. I argue that trading offers cost savings and efficiency in production when firms have economies of scale. That is, if the average cost of a firm is decreasing, trading will allow the firm to expand their operation and produce at lower per-unit costs.

To identify the impacts of trading caps on market power and production costs, I use the difference-in-difference strategy. I exploit the fact that only a subset of vessels in Norwegian cod fishery is allowed to trade quotas because they have tradable licenses, and this tradability status is defined on the vessel length three years before the trade program was implemented (for the precedent non-tradable cap implementation rather than an anticipation for CAT). Hence, my difference-in-difference estimates compare the changes in the outcomes of interest from before the trade program year to after the policy program between vessels in the trade qualified group and vessels in the unqualified group to identify the causal effects of the quota trading program.

I first apply the difference-in-difference strategy to examine the impacts of trading and quota acquisition on standard fishing outcomes, namely harvest quantity and revenue. I then test the market power hypothesis by looking at the impact on observed transacted fish sales prices when a vessel lands their catch. The cost efficiency hypothesis and the mark-up examination require studying the change in harvest costs, which is not observed. I overcome this challenge by introducing a method to estimate the output elasticity of total cost using production function estimation and data on just output and input quantity. I show how the cost elasticity can infer the shape of production costs and economies of scale. This method relies on just the standard cost minimizing condition of the input choice problem, and is applicable in a general context, not just fishery, where data on output and input are available and allow us to estimate a production function.

Results show the CAT program has dramatically increased the harvest quantity and revenue of a vessel. There are cases where vessels acquire additional quotas from the trading program double their harvest and revenues. This is apparent given the anticipation of consolidation in the trading scenario. Given the intense consolidation in the tradable groups, vessels staying in these groups have sold their fish at a little higher transacted prices than before (by 2% or 0.43 Norwegian krone per kg, equivalent to 2 US cents per lb). However, evidence on the change in price is noisy and the market concentration Herfindahl-Hirschman Index has been still very low even after the trading program, that is, below 0.03. Hence, in general, I conclude the trading program does not lead to an increase in the sales prices in the Norwegian fish market.

The rejection of the change in prices drives an attention to the cost efficiency of the CAT policy. The reason is vessels that acquire additional quotas from the trading program must be going to have their operation unit cost lower than before so that they have incentives to expand the harvest and gain profit margins to cover the quota expenses. My finding confirms this hypothesis of cost advantages. First, I find that the output elasticity of cost is *less* than 1 for all vessels *before* the trading program. Because the output elasticity of cost is the ratio of marginal cost to average cost, the inelastic measure suggests vessels were operating at which marginal cost is *below* average cost. In other words, they are on the decreasing side of the average cost curve. Second, the difference-in-difference estimator shows trading policy has significantly increased the cost elasticity. Intuitively, trading has raised the output elasticity of cost toward 1. These two findings suggest the CAT has moved vessels' operations toward the minimum of average cost where average cost equals marginal cost. Final calculation finds a vessel in the largest licensed length group (21–28m) on average acquires caps that increased catch by 50% and decreased average costs per tonne by 15.60%. Given no change in prices, the finding of cost reduction suggests vessels' owners have implicitly obtained significant market power after trading in which they could have lowered prices but did not.

As mentioned, I offer a new perspective of CAT: cost efficiency owing to economies of scale. This adds to the literature of CAT that has shown CAT offers cost effectiveness due to productivity heterogeneity. I propose a method to decompose the output growth into the cost-efficiency component and the cost-effectiveness component. I show that the cost efficiency captures the output change owing to economies of scale and the movement of the operation along an average cost curve, whereas the cost effectiveness captures the output change due to productivity improvement and the shifts of the cost curve. In the studied Norwegian fishery context, the contribution of cost efficiency to the output growth varies substantially with the size of license (hence quotas). For the smallest tradable vessel group, the cost efficiency does not contribute at all to the change in output. However, it accounts for more than 50% (and even 100% in the group of biggest vessels) of the output growth in the other bigger groups in the first three years since the first time of acquiring quotas. The cost effectiveness owing to productivity improvement contributes most to the output growth in subsequent years.

In summary, this paper strives for a two-fold goal to not only look at the causal effects of cap-and-trade (CAT) relative to command-and-control (CAC) on the firm performance in the product market in resource economics, and also to offer a method to estimate economies of scale that is applicable in other empirical studies in economics.

In environmental economics, the relative ex-post performance between CAT and CAC has been difficult to test empirically due to challenges in constructing a credible benchmark we would have observed in the non-trading program to identify the causal effect of cap-and-trade. An exception is Fowlie, Holland and Mansur (2012) that exploit the variation in the participation requirements of the RECLAIM program and uses matching difference-in-difference estimator to identify the causal effects of cap-and-trade on emissions and the distributional effects by facility neighborhood demographic characteristics. Petrick and Wagner (2014) study the causal effect of CAT on input reallocation in manufacturing sectors. They find CAT firms curb the consumption of natural gas and petroleum products but not electricity use and no evidence on lower employment, gross output, or exports. Another branch in this literature is to use structural models to estimate abatement costs of CAT relative to CAC; see Carlson et al. (2000); Chan et al. (2018). This paper focuses on the ex-post impact of CAT on firm performance in the main product market rather than by-products such as emissions.

Investigating the firm performance in the main product market is even more important in the fishery context, because the cap regulation in fishery directly imposes a constraint on the main product rather than by-products. CAT in fishery, hence, likely incentivizes a firm to allocate production inputs and exploit economies of scale to expand output at lower average cost of a unit. Trading in fishing quotas, by allowing for the consolidation of quotas and output, also raises strong concerns about market power abuse. This paper is the first empirical research that identifies the causal effect of tradable quotas on the firm performance and the implications of trading on market power in fishery. Previous studies attempted to analyze the firm performance using the pre-policy and post-policy data but failed to control counterfactual trend in the absence of tradable quota policy to identify the causal relation; see Grafton, Squires and Fox (2000); Fox et al. (2003). Some recent studies have estimated the causal effects of tradable quotas using program evaluation designs, but they have investigated the impacts on stock biomass indices and probability of a fish stock collapsing using a global database of fisheries institutions and catch statistics; see Costello, Gaines and Lynham (2008); Costello et al. (2010); Isaksen and Richter (2019).

This paper is also the first study that brings the production function estimation in the IO and macroeconomics literature to environmental and resource economics. Estimating a production function has been useful and important for two reasons. First, one can obtain the estimate of productivity TFPQ and then investigate the dynamics of TFPQ over time and across firms. A few studies recently have combined the production function estimation with a difference-in-difference framework to study the impact of industrial events such as exporting status and mergers and acquisitions; see De Loecker (2013); Braguinsky et al. (2015); Stiebale and Vencappa (2018); Rubens (2021). This paper applies the production function estimation to study the impact of cap-and-trade in fishery.

The second advantage of the production function estimation is one can even recover the markup (the price-cost ratio) from the estimates of a production function. This approach to recover markup has been so-called the production approach to distinguish it from the demand approach that estimates the demand and relies on assumptions on how firms compete in the market. Examples of the demand approach are huge, following advances in demand estimation by Berry (1994) and Berry, Levinsohn and Pakes (1995). In contrast to the demand approach, the production approach relies on the classic cost-minimizing behavior in firms' input allocation and the observed input's expenditure share in revenue in most of financial statement reports, following the work by De Loecker and Warzynski (2012). In this paper, I unfortunately do not observe the share of input cost in revenue to explore the change in markup.

However, and most importantly, I offer a method to measure the economies of scale from the estimates of the production function and to infer changes in average costs.[1] I show that economies of scale can be measured by either the output elasticity of total cost or the output elasticity of average cost. In fact, the two elasticities differ just by one percentage point, and the output elasticity of total cost is the ratio of marginal cost to average cost. With the additional assumption of the classic cost minimizing behavior, I show the output elasticity of total cost is the reciprocal of the sum of all output elasticities of inputs. Hence, one can estimate economies of scale using the output elasticities of inputs that are implied from the estimates of the production function, and this method is applicable beyond the context of environmental and resource economics.

Besides using the measure of cost elasticity to infer economies of scale, I discuss other applications of the measure: inferring markups if the cost revenue ratio is observed, testing the validity of production function estimates, and decomposing output growth into factors due to economies of scale vis à vis productivity differences. I show that the decomposition has three interpretations. From the perspective of production, the decomposition decomposes output growth into change due to input reallocation and change due to productivity improvement. From the perspective of production costs, the decomposition breaks down the output change into the change due to sliding on the average cost curve and the change due to shifting the cost curve. From the perspective of quota trading, the input-adjustment effect captures the output change due to economies of scale or cost efficiency, whereas the productivity-change effect captures the output change due to cost effectiveness. Hence, without knowing the information of quota acquirees, we can still decompose the output growth into the cost-efficiency effect and the cost-effectiveness effect. This contributes to the decomposition analyses that are used to analyze the relative importance of various factors; see the decomposition techniques in other contexts by Holland et al. (2020); De Loecker, Eeckhout and Unger (2020); Haltiwanger (1997); Fortin, Lemieux and Firpo (2011); Ang and Zhang (2000).

This paper is outlined as follows. Section 2 presents the theory of estimating economies

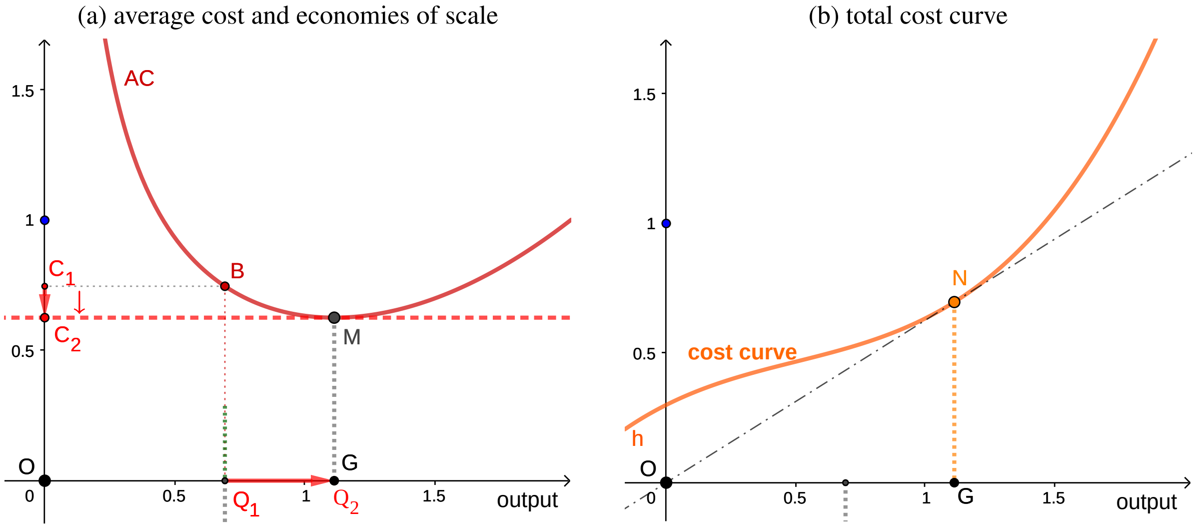
---

[1] Together with the observed change in observed transacted fish sales prices, I can infer the change in markup.

of scale using data on output and input quantity and applications of the economies-of-scale measure. Section 3 summarizes the Norwegian cod fisheries regulations. Section 4 presents the empirical strategies that use difference-in-difference (DID) methodology and instrumental-variable difference-in-difference (IV DID) methodology to estimate the intent to treat and local average treatment effect of cap-and-trade program. The section also discusses the approach to estimate a production function. Section 5 describes data sources and summary statistics. Section 6 provides results. Section 7 discusses the decomposition of output into the component due to economies of scale and the one due to cost shifting. Section 8 concludes.

2 Economies of Scale and Applications

2.1 Measuring


(a) average cost and economies of scale

(b) total cost curve

Figure 1: The theory of production cost and economies of scale

Economies of scale refer to the cost advantages that a firm obtains for its scale of operation. Consider the average cost in Figure 1a and its originating cost curve in Figure 1b. The movement from B to M exhibits economies of scale: As quantity of production increases from Q_1 to Q_2, the average cost of each unit decreases from C_1 to C_2. Formally, economies of scale are the output production segment in which the average cost is decreasing.

On the other hand, diseconomies of scale are the output segment in which the average cost is increasing. Let M be the minimum average cost point, which corresponds to point N on the total cost curve where its average cost equals the slope of the cost curve. In Figure 1a, all points on the left of the minimal average cost M exhibit economies of scale because an increase in output is associated with a decrease in average cost. On the other hand, points on the right of M illustrate diseconomies of scale because their average costs are increasing in the firm's size.

To measure economies of scale, I use the output elasticity of average cost that measures the percentage change in average cost when output increases by one percent: $\psi \equiv \frac{dAC}{dq} \cdot \frac{q}{AC}$. Negative elasticities are equivalent to economies of scale, whereas positive elasticities mean diseconomies of scale.

Another useful measure is the output elasticity of total cost that gives the percentage change in total cost when output increases by one percent: $\phi \equiv \frac{dC}{dq} \cdot \frac{q}{C}$. This measure is effectively the ratio of marginal cost to average cost. For every differentiable cost function, the two elasticities differ just by one unit:

$$\psi \equiv \frac{dAC}{dq} \cdot \frac{q}{AC} = \left(\frac{C(q)}{q}\right)' \cdot \frac{q}{AC} = \frac{MC - AC}{AC} = \phi - 1, \tag{1}$$

where AC denotes average cost of each unit and MC denotes marginal cost. Proposition 1 below summarizes these results.

Proposition 1. *Assume the cost function is first differentiable. Let ψ and ϕ denote output elasticities of average cost and total cost, respectively. We have the following results:*

i) $\psi = \phi - 1$.

ii) *ϕ is the ratio of marginal cost to average cost. ψ is the percentage difference between marginal cost and average cost.*

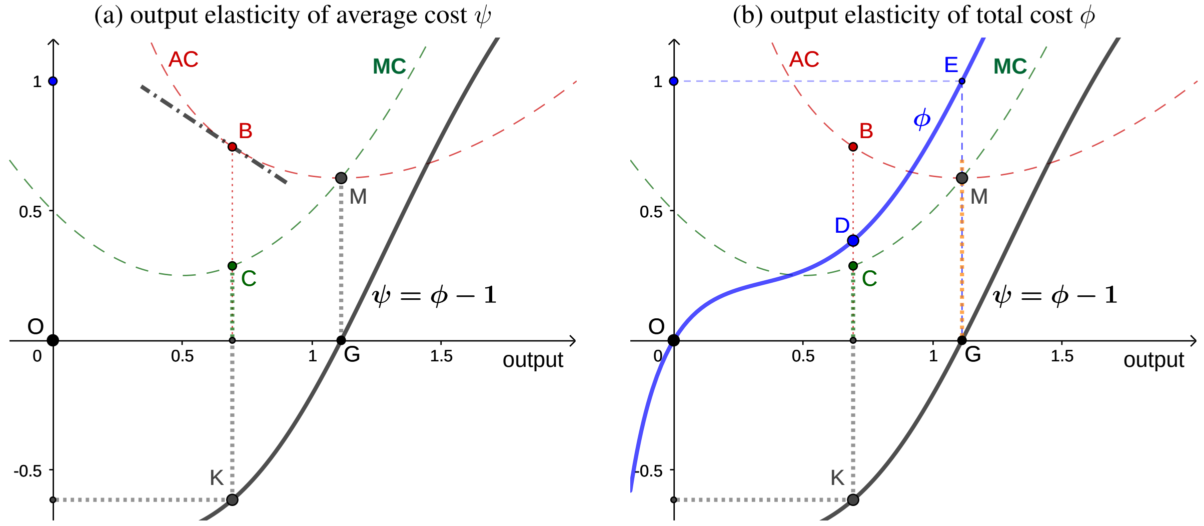
iii) *An economy of scale (decreasing average cost) is equivalent to $\psi < 0$ or $\phi < 1$. A diseconomy of scale is equivalent to $\psi > 0$ or $\phi > 1$. A constant economy of scale means $\psi = 0$ or $\phi = 1$.*

Note that the results apply for every first differentiable cost function, regardless whether the cost is (locally) convex. The reason is the sign of ψ equals the sign of the average cost by its definition. If the cost function is convex, then a point of constant economy of scale is a minimum of the average cost level. The reason is the convexity (the sign of the second derivative) of the average cost equals the convexity of the total cost at a point where marginal cost equals average cost.

Figure 2 illustrates the relations between average cost, marginal cost, and output elasticities of average cost and total cost. For output levels on the left of point G, their average costs are above marginal cost and decreasing. These output levels exhibit economies of scale and have $\psi < 0$ or $\phi < 1$. Output levels on the right of G exhibit diseconomies of scale with $\psi > 0$, $\phi > 1$, marginal cost above average cost, or increasing average cost. Assuming the cost function is convex, then the constant economy-of-scale point G with $\psi = 0$ ($\phi = 1$) is the output of the minimum of average cost M.

2.2 Estimating Using Production Data

As shown, the two measures ψ and ϕ have one-to-one relation and one just needs to estimate either of them. I now discuss the method to estimate ϕ using data on output and input quantity.


(a) output elasticity of average cost ψ

(b) output elasticity of total cost ϕ

Figure 2: Measuring economies of scale by output elasticities of average-cost and total costs

The method relies on the standard definition of a long run cost. Specifically, consider the cost-minimization problem of firm i to produce the targeted output q_{it} at time t:

$$C_{it}(q_{it}) = \min_{\boldsymbol{X}_{it}} \boldsymbol{W}_{it}^{\top}\boldsymbol{X}_{it} \text{ subject to } q_{it} \leq Q_{it}(\boldsymbol{X}_{it}), \tag{2}$$

where $\boldsymbol{X}_{it}$ is the input vector, $\boldsymbol{W}_{it}$ is the input-price vector, and $Q_{it}(\boldsymbol{X})$ is the production technology. Assume the production function $Q_{it}(\cdot)$ is continuous and twice differentiable. The Lagrangian function of the cost-minimization problem is

$$\mathcal{L}_{it} = \boldsymbol{W}_{it}^{\top}\boldsymbol{X}_{it} + \lambda_{it}(q - Q_{it}(\boldsymbol{X}_{it})). \tag{3}$$

The first-order condition for any input $X \in \mathbb{X}$ is

$$\frac{\partial \mathcal{L}_{it}}{\partial X_{it}} = W_{it} - \lambda_{it} \cdot \frac{\partial Q_{it}}{\partial X_{it}} = 0. \tag{4}$$

Rearranging terms and multiplying by $\frac{X_{it}}{Q_{it}}$, we get

$$\frac{W_{it}X_{it}}{Q_{it}} \cdot \frac{1}{\lambda_{it}} = \frac{\partial Q_{it}}{\partial X_{it}} \cdot \frac{X_{it}}{Q_{it}}, \tag{5}$$

Notice that equation (5) applies for every input $X \in \mathbb{X}$. Hence, I sum all these relations side by side to get

$$\frac{\sum_{X \in \mathbb{X}}(W_{it}X_{it})}{Q_{it}} \cdot \frac{1}{\lambda_{it}} = \sum_{X \in \mathbb{X}} \left(\frac{\partial Q_{it}}{\partial X_{it}} \cdot \frac{X_{it}}{Q_{it}} \right). \tag{6}$$

Note that $\sum_{X\in\mathbb{X}}(W_{it}X_{it})$ is total cost, and λ is marginal cost of production because $\frac{dC_{it}}{dq} = \frac{\partial\mathcal{L}^*_{it}}{\partial q} = \lambda_{it}$. Hence, the left hand side is the ratio of average cost to marginal cost, or reciprocal ϕ. The right hand side is the sum of all input elasticities of output. Denote the input X elasticity of output as θ^X_{it}, $\theta^X_{it} \equiv \frac{\partial Q_{it}}{\partial X_{it}} \cdot \frac{X_{it}}{Q_{it}}$. I can conclude the following proposition.

Proposition 2 (Cost elasticity and input elasticity)**.** *Assume cost-minimization behavior in a price-taker input market. Then, output elasticity of total cost is the reciprocal of the sum of all input elasticities of output:*

$$\phi_{it} = \left(\sum_{X\in\mathbb{X}} \theta^X_{it}\right)^{-1}.$$

Alternatively, the ratio of average cost to marginal cost is the sum of all input elasticities of output.

Proposition 2 implies that we can use data on output and input quantity to estimate the production function, derive the input elasticities of output, and calculate the output elasticity of total cost. In the empirical part of this paper, I estimate the production function using OLS with fixed effects, the proxy variable method (Ackerberg, Caves and Frazer, 2015), and the dynamic panel approach (Blundell and Bond, 2000). Each method of estimating a production function relies on different identifying assumptions. Section 4.2 discusses their identifying assumptions and the compatibility between them and the assumptions underlying the cost elasticity formula.

The assumptions underlying the cost elasticity formula, as we have seen, rely only on the assumption of cost minimization in the input choice problem.

Assumption 1. Proposition 2 relies on the following assumptions:
(1.1) Input prices are exogenous, that is, they do not depend on input quantities.
(1.2) All inputs are variable.

Assumption (1.1) implies firms are price-takers in the input market. That is, variation in input prices comes from exogenous factors rather than firms' input usage. If the firm's input quantity affects the input price, the relation between output elasticity of cost and input elasticities of output involves the price elasticity of input demand. Appendix B derives this relation. In that case, one would need additional information on the price elasticity of input demand to estimate ϕ.

Assumption (1.2) implies the cost function is the long run cost, because all inputs are variable. If there is a fixed input, then the formula in Proposition 2 excludes the consideration of the fixed input and measures the output elasticity of total *variable* cost and the reciprocal of the sum of all *variable* inputs elasticities of output.

2.3 The Case of Dynamic Inputs

This extension quantifies the cost elasticity when an input has adjustment costs and dynamic implications on future cost values. Consider the classical cost minimization in a dynamic context in which capital K_{it} is dynamic and adjusted by endogenous investment level $I_{i,t-1}$ whereas labor L_{it} is variable. So, the capital evolves as $K_{it} = \delta K_{i,t-1} + I_{i,t-1}$ and the adjustment cost depends on both investment level and the capital state, $A(I_{i,t-1}, K_{i,t-1})$. For simplicity, I drop the notation i in this section. The dynamic cost minimization problem is

$$\begin{aligned} V(K_{t-1}, \Omega_t) &= \min_{I_{t-1}, L_t} rI_{t-1} + wL_t + A(I_{t-1}, K_{t-1}) + \beta E[V(K_t, \Omega_{t+1})|\Omega_t], && (7)\\ \text{subject to } & Q(K_t, L_t) \geq q_t, && (8)\\ & K_t = \delta K_{t-1} + I_{t-1}. && (9) \end{aligned}$$

Note that we can rewrite this problem into

$$V(K_{t-1}, \Omega_t) = \min_{I_{t-1}, L_t} r \cdot (K_t - \delta K_{t-1}) + w \cdot L_t + A(K_t - \delta K_{t-1}, K_{t-1}) + \beta E[V(K_t, \Omega_{t+1})|\Omega_t].$$

So, we can consider an equivalent problem with endogenous choices of capital and labor:

$$\begin{aligned} V(K_{t-1}, \Omega_t) &= \min_{K_t, L_t} rK_t + wL_t + \mathcal{A}(K_t, K_{t-1}) + \beta E[V(K_t, \Omega_{t+1})|\Omega_t],\\ \text{subject to } & Q(K_t, L_t) \geq q_t. && (10) \end{aligned}$$

Proposition 3 (The dynamic version of the ratio of AC to MC). *In a dynamic cost minimization with adjustment costs such that* $\mathcal{A}(K^*_{t+1}, K^*_t) = K^*_t \cdot \frac{\partial \mathcal{A}}{\partial K_{t+1}} + K^*_t \cdot \frac{\partial \mathcal{A}}{\partial K_t}$, *we have*

$$\begin{aligned} \frac{AVC_t + E[AAC_{t+1}|\Omega_t]}{MC_t} &= \theta_{L_t} + \theta_{K_t},\\ \textit{where } AVC_t &= \frac{rK_t + wL_t}{Q_t},\\ AAC_{t+1} &= \frac{\mathcal{A}(K_{t+1}, K_t)}{Q_t}. \end{aligned}$$

Note that $AVC_t = E[AVC_t|\Omega_t]$ and $MC_t = E[MC_t|\Omega_t]$. Intuitively, we have a dynamic equivalent version for the Proposition 2: The ratio of expected average cost to marginal cost, as defined by the ratio of total variable cost and expected adjustment cost to marginal cost, is the sum of all input elasticities of output.

Proof. Consider the dynamic problem (10), the FOCs with respect to L_t, K_t are:

$$w = \lambda \cdot \frac{\partial Q}{\partial L_t} \Longrightarrow \frac{w \cdot L_t}{\lambda \cdot Q} = \frac{\partial Q}{\partial L_t} \cdot \frac{L_t}{Q} = \theta_{L_t} \tag{11}$$

$$r + \frac{\partial \mathcal{A}(K_t, K_{t-1})}{\partial K_t} + \beta E\left[\frac{\partial V(K_t, \Omega_{t+1})}{\partial K_t} | \Omega_t\right] = \lambda \cdot \frac{\partial Q}{\partial K_t} \tag{12}$$

To rewrite the FOC wrt K_t, firstly use the Envelope theorem to calculate the derivative of value function:

$$\frac{\partial V(K_t, \Omega_{t+1})}{\partial K_t} = \frac{\partial \mathcal{A}(K_{t+1}, K_t)}{\partial K_t}. \tag{13}$$

Substitute this in the FOC wrt K_t to get the Euler equation for the dynamic capital:

$$r + \frac{\partial \mathcal{A}(K_t, K_{t-1})}{\partial K_t} + \beta E\left[\frac{\partial \mathcal{A}(K_{t+1}, K_t)}{\partial K_t} | \Omega_t\right] = \lambda \cdot \frac{\partial Q}{\partial K_t} \tag{14}$$

Multiply both sides by $\frac{K_t}{Q(K_t, L_t) \cdot \lambda}$ at optimal levels K_t^*:

$$\Longrightarrow K_t^* \cdot \left(r + \frac{\partial \mathcal{A}(K_t, K_{t-1})}{\partial K_t} + \beta E\left[\frac{\partial \mathcal{A}(K_{t+1}, K_t)}{\partial K_t} | \Omega_t\right]\right) / (Q \cdot \lambda) = \frac{\partial Q}{\partial K_t} \cdot \frac{K_t}{Q} \equiv \theta_{K_t}|_{K_t^*},$$

$$\Longrightarrow \frac{K_t^* \cdot r + K_t^* \cdot \mathcal{A}_1 + \beta \cdot K_t^* \cdot E[\mathcal{A}_2 | \Omega_t]}{Q \cdot \lambda} = \theta_{K_t}|_{K_t^*}, \tag{15}$$

where $\mathcal{A}_1$ and $\mathcal{A}_2$ denote the first and second derivatives of $\mathcal{A}$.

If the adjustment cost satisfies $\mathcal{A}(K_{t+1}^*, K_t^*) = K_t^* \cdot \mathcal{A}_1 + \beta \cdot K_t^* \cdot \mathcal{A}_2$, then

$$\frac{E[(K_t^* \cdot r + \mathcal{A}(K_{t+1}, K_t)) | \Omega_t]}{Q} = \theta_{K_t}|_{K_t^*}. \tag{16}$$

Then sum the equalities (11) and (16) side by side, we get

$$E\left[\frac{AVC + AAC}{MC}\right] = \theta_{L_t} + \theta_{K_t}. \tag{17}$$

□

2.4 Compatibility of Profit Maximization and Cost Minimization

We have seen that the relation between output elasticity of cost and input elasticity of output relies on the cost-minimization behavior of producers. Traditionally, we often see the goal of the firm is to maximize profit. So, one may wonder whether the two behaviors are incompatible. I now show that the two behaviors are compatible in a variety of market structures: perfect competition, Cournot competition, Cournot competition in the presence of bargaining power stemming from output size, price differentiation due to output-independent quality adjustment, and co-influence of output and input in a generalized cost function.

Proposition 4 (Compatibility of profit maximization and cost minimization)**.** *Assume differentiability, concavity of profit function, and convexity of cost function. Let the input choice problem to maximize profit be*

$$\textit{[Problem 1:]} \quad \max_{\boldsymbol{X}_{it}} \mathcal{P}_{it}(Q_{it}(\boldsymbol{X}_{it})) \cdot Q_{it}(\boldsymbol{X}_{it}) - \mathcal{G}(\boldsymbol{X}_{it}),$$

where $\mathcal{P}_{it}(\cdot)$ is the firm individual output price function, $Q_{it}(\boldsymbol{X}_{it})$ is the production function, $\mathcal{G}(\cdot)$ is the generalized cost function. Let the two-step decision problem where the firm decides output level to maximize profits in the first stage and decides inputs to minimize production cost of producing the targeted output in the second stage be

$$\begin{aligned}\textit{[Problem 2:]} \quad & \max_{q_{it}} \ \mathcal{P}_{it}(q_{it}) \cdot q_{it} - C(q_{it}) \textit{ in stage 1, and} \\ & C(q_{it}) = \min_{\boldsymbol{X}_{it}} \ \mathcal{G}(\boldsymbol{X}_{it}) \textit{ subject to } Q_{it}(\boldsymbol{X}_{it}) \geq q_{it} \textit{ in stage 2.}\end{aligned}$$

The two problems are equivalent for the following market environments.

i) Perfect competition. That is, firms are price takers in the output market.

ii) Cournot competition. That is, individual firm price is the common market price: $\mathcal{P}_{it} = P(\mathcal{Q}(q_{it}))$, where $\mathcal{Q}$ is the total output of all firms in the market.

iii) Bargaining power stemming from output size. That is, $\mathcal{P}_{it} = P(\mathcal{Q}(q_{it}), q_{it})$.

iv) Price differentiation stemming from endogenous efforts that affect quality but not quantity. That is, $\mathcal{P}_{it} = P(\mathcal{Q}(q_{it}), q_{it}, H(e_{it}))$ and $\mathcal{G}(\cdot) = \mathcal{G}(\boldsymbol{X}_{it}, e_{it})$.

v) Co-influence of output and input in cost function. That is, $\mathcal{G}(\cdot) = G(Q_{it}(\boldsymbol{X}_{it}), \boldsymbol{X}_{it})$.

If firms can differentiate their prices by allocating inputs to directly adjust product quality $H(\boldsymbol{X}_{it})$, i.e. $\mathcal{P}_{it} = P(\mathcal{Q}(q_{it}), q_{it}, H(\boldsymbol{X}_{it}))$, then the two problems are not equivalent in general.

Appendix C shows the proof. The main intuition is that the cost minimization problem lies in the production stage rather than being a whole single goal of the firm. The cost minimization problem aims to design inputs to produce the targeted output rather than to design the output that minimizes cost.

Proposition 4 generates an important implication on the connection between the formula of cost elasticity in this paper and the methodologies of estimating a production function in the literature. The fact that the formula of cost elasticity does not rely on the perfectly competitive environment allows us to use a variety of methodologies of estimating a production function to empirically estimate economies of scale.

In later sections where I apply the theory of economies of scale in the fishery context, I assume the price differentiation, if any, comes from the bargaining power owing to output size or from the fish quality affected by exogenous factors and/or endogenous factors beyond production inputs. Endogenous factors beyond production inputs are endogenous factors a fisherman can control but do not enter directly in the harvest process. An example is the on-board processing and freezing storage that improve the fish quality but they do not affect the harvest quantity. Another example is the vessel engine power that helps speed up the return trip to quickly reach the sale points in mainland and keep the fish freshness. With this assumption, the input choice to maximize profits will be equivalent to the 2-step decision process where the firm adjusts inputs to minimize the production costs of a targeted output level.

Before moving to the specific application of economies of scale in the fishery cap-and-trade, I conclude Section 2 by discussing three applications of economies of scale in a general context.

2.5 Application 1: Estimating Markups

Although I am the first that derives elasticities of costs to measure economies of scale and shows how to estimate it, I am not the first that exploits the cost-minimization condition of the input allocation problem. Indeed, an emerging literature on IO and macroeconomics has used this condition to estimate markups. This approach has been called production approach to distinguish it from the demand approach. In the demand approach championed by Bresnahan (1989) and Berry, Levinsohn and Pakes (1995), the markup estimation relies on assumptions on utility maximizing behavior of consumers and on how firms compete (for example, Bertrand-Nash price competition or Cournot quantity competition). This demand approach requires data on (at least) product market shares and product characteristics. In contrast, the production approach, established by De Loecker and Warzynski (2012), is posited on the cost minimization by producers and requires data on individual firm output, input, and a variable input's expenditure share in revenue. Examples of applications of this production approach include Braguinsky et al. (2015); De Loecker et al. (2016); De Loecker, Eeckhout and Unger (2020). I now discuss the relation between my output elasticity of cost and the markup of this production approach literature.

Let me begin with the review of the production approach. The production approach to estimate markups also relies on the cost minimizing problem (2). However, De Loecker and Warzynski (2012) rewrite equation (5) into

$$\frac{W_{it}X_{it}}{P_{it}Q_{it}} \cdot \frac{P_{it}}{\lambda_{it}} = \frac{\partial Q_{it}}{\partial X_{it}} \cdot \frac{X_{it}}{Q_{it}}, \tag{18}$$

where P_{it} is the output price. Hence, the markup ratio $\mu \equiv \frac{P_{it}}{\lambda_{it}}$ can be calculated through:

$$\mu_{it} = \frac{\theta^X_{it}}{\alpha^X_{it}}, \tag{19}$$

where α^X_{it} is the share of expenditures on input X in total sales, i.e. $\alpha^X_{it} \equiv \frac{W_{it}X_{it}}{P_{it}Q_{it}}$. Using this relation, De Loecker and Warzynski (2012) show firm-level markups can be inferred using production data. Specifically, one would need (i) data on output and input to estimate the production function and the output elasticity of one (or more) variable input(s) θ_{it}^X and (ii) data on expenditure share α_{it}^X, which is often available in the financial statement of the firms.

To think about the relation between output elasticity of cost ϕ and markup μ, we now can use Proposition 2 and equation (19) to get

$$\sum_{X\in\mathbb{X}} \theta_{it}^X = \sum_{X\in\mathbb{X}} \mu_{it}\alpha_{it}^X = \mu_{it}\sum_{X\in\mathbb{X}} \alpha_{it}^X = \mu_{it}\cdot\frac{C_{it}(Q_{it})}{P_{it}Q_{it}} \tag{20}$$

$$\implies \phi_{it}\cdot\mu_{it} = \frac{PQ}{C} \tag{21}$$

However, I want to state this relation in a separate proposition, because this relation in fact exists *without* the cost minimization condition.

Proposition 5 (Cost elasticity and markup)**.** *Assume the cost function is differentiable and denote the markup as $\mu \equiv \frac{P}{MC}$. Then, we have the following relation between output elasticity of cost ϕ and markup μ:*

$$\phi\cdot\mu = \frac{PQ}{C} \equiv \frac{revenue}{cost} \tag{22}$$

Proof. $\phi\cdot\mu = \frac{dC}{dQ}\cdot\frac{Q}{C}\cdot\frac{P}{dC/dQ} = \frac{PQ}{C}$ □

Again, I want to emphasize that the result in Proposition 5 requires only one condition: Cost is differentiable. Hence, it is the relation between ϕ and θ or between μ and θ that requires the cost minimizing behavior in the input choice decision.

With this result, we now have two ways to estimate markup in the family of the production approach. One is to follow De Loecker and Warzynski (2012): Using equation (19) if the share of expenditure on a specific input X in revenue, α^X is observed. The other one is to follow my Propositions 2 and 5 if the cost revenue ratio is observed.

2.6 Application 2: Testing the Estimate of a Production Function

Propositions 1 and 2 imply we can use the cost elasticity (as the reciprocal of total input elasticities of output) as a measure to test the reasonability of the estimate of a production function. Specifically, a plausible estimate of a production function should result in plausible estimates of output elasticities of inputs θ and thus, ϕ. Because ϕ is the ratio of marginal cost to average cost, a reasonable estimate of ϕ should be not much below 0 and not much above 1. A big negative estimate of ϕ is hard to convince because it implies not only a high average cost, although the firm is having an economy of scale, but also a negative marginal cost. An estimate of ϕ that is hugely higher than 1 is also difficult to justify, because $\phi > 1$ implies the firm has passed by its economy of scale and is producing at which marginal cost is above average cost. $\phi \times 100$ is percentage of marginal cost in average cost. Hence, a huge estimate of ϕ, say 100, would be implausible in many industries. If we know the rule of thumb for the practical range of the output elasticity of cost (or average cost) in an industry, we can have a rule of thumb for the production function estimates.

Although assessing the credibility of the estimate of ϕ and a production function is possible, investigating the exact reason of the bias is difficult. There can be several reasons for the bias. First, a production function (and hence cost function) might be misspecified by under-including or over-including a production factor. For example, consider the case where the production function misses one input. If marginal cost is positive and although the estimates of output elasticities of other inputs are not biased, then we would miss a positive output elasticity of the missing input when calculating the sum of all necessary output elasticities of inputs. Then, we would over-estimate ϕ, if we under-include an input.

However, if the under-inclusion of an input results in biased estimates of output elasticities of other inputs in the estimation stage, the bias of the estimated $\widehat{\phi}$ accommodates both bias due to misspecification (upward bias) and bias due to confounded $\widehat{\theta}$. The bias, hence, may be either downward or upward, depending on the sign of bias in the confounded $\widehat{\theta}$ and the competing direction to the misspecification.

Another reason for the bias in estimated $\widehat{\phi}$ is the misspecified cost function. As shown in Appendix B, if input prices are not exogenous to firms' input choices, then ϕ may be marked down or marked up, depending on whether the price elasticities of each input demand are inelastic or elastic.

2.7 Application 3: Decomposing Output Growth into Economies-of-Scale Factors and Cost-Shifting Factors

One contribution of this paper is to illustrate the cost efficiency vis à vis cost effectiveness of cap-and-trade in fishery. The literature on cap and trade has shown that trading offers cost effectiveness in which the cap can be transferred from low productive (or high cost) firms to high productive (or low cost) firms, effectively equalizing the marginal value of cap across firms and achieving the second-best solution of social welfare maximization. The cost efficiency perspective in this paper refers to the within-firm cost savings due to the firm's input reallocation in response to an event (trading or a policy). Formally, the cost effectiveness is implied by the change in productivity whereas the cost efficiency is implied by the change in input reallocation. I now propose a decomposition method to see how much of the output change is attributed to input reallocation and to productivity change.

Note that the case of cap-and-trade is just an example. In general, a policy or an event may affect both input allocation and productivity. I show that the decomposition has three interpretations. From the perspective of production, it decomposes the output change into the change due to input adjustment and the change due to productivity change. From the perspective of operation costs, it breaks down output change into the change owing to economies of scale (or the slide along the average cost curve) and the change owing to shifting the cost curve. From the perspective of trade, it illustrates the output change due to cost efficiency (within-firm economies of scale) and cost effectiveness (heterogeneity of cost or productivity in comparison with trade partners).

Proposition 6. *Assume the production function $Q_{it} = \mathbb{F}(\mathbf{X}_{it}, \omega_{it})$ contains only Hicksian productivity. The dynamic change in output can be decomposed into the output change due to economies of scale and due to shifting the cost curve. That is,*

$$\Delta Q_{it} \equiv Q_{it} - Q_{i,t-1} = \underbrace{\overbrace{\mathbb{F}(\mathbf{X}_{it}, \omega_{it}) - \mathbb{F}(\mathbf{X}_{i,t-1}, \omega_{it})}^{\textit{input reallocation } \Delta Q|_{\Delta X}}}_{\textit{economies of scale}} + \underbrace{\overbrace{\mathbb{F}(\mathbf{X}_{i,t-1}, \omega_{it}) - \mathbb{F}(\mathbf{X}_{i,t-1}, \omega_{i,t-1})}^{\textit{productivity change } \Delta Q|_{\Delta \omega}}}_{\textit{cost shifting}}.$$

Whereas the decomposition is straightforward, Appendix D formally proves that the change in output due to input reallocation refers to the change in output along an average cost curve, whereas the change in output due to productivity change is owing to cost shifting.

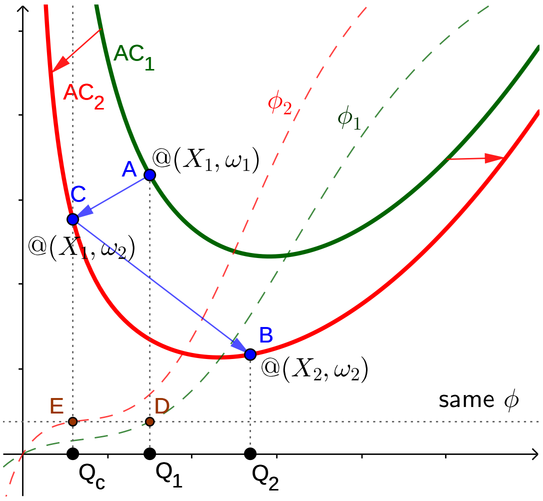


Figure 3: Decomposition of output change: Cost shifting due to productivity change and cost moving due to economies of scale

Note: Moving from A to C depicts a change due to cost shifting, which is attributed to heterogeneous productivity while keeping the cost elasticity ϕ unchanged. Moving from C to B depicts a change due to economies of scale. Figure H15 in the appendix illustrate other scenarios of cost shifting.

Intuitively, the cost elasticity, a measure of economies of scale, is a sufficient statistics for the change in all inputs. Figure 3 graphically illustrates the decomposition. Suppose that in period 1, the firm produces output Q_1 using input bundle $\mathbf{X}_1$ at productivity level ω_1. This output level corresponds to the cost level A on the average cost curve AC_1 (the green solid curve) and the output elasticity of total cost D on the cost elasticity curve ϕ_1 (the green dash curve). Note that the average cost curve AC_1 and cost elasticity curve ϕ_1 are the cost-related functions corresponding to the productivity level ω_1. Any change in the input bundle while keeping the productivity ω_1 constant is the slide along the curves AC_1 and ϕ_1.

In period 2, for any reason, the firm changes its input bundle from $\mathbf{X}_1$ to $\mathbf{X}_2$ and productivity from ω_1 to ω_2. With the new productivity level ω_2, the new cost-related functions are the average cost curve AC_2 (the red solid curve) and the cost elasticity ϕ_2 (the red dash curve). Let Q_2 be the firm's new output, which is point B on the average cost curve AC_2 that corresponds to input bundle $\mathbf{X}_2$ and productivity ω_2.

We are interested in the decomposition of output and cost from A to B, that is, from production level Q_1 ($Q_1 = \mathbb{F}(\mathbf{X}_1, \omega_1)$) to Q_2 ($Q_2 = \mathbb{F}(\mathbf{X}_2, \omega_2)$). Graphically, the change is a result of the cost shifting from point A to point C due to productivity change and the cost sliding from point C to point B due to input adjustment. The intermediary point C represents the average cost level on the second-period cost curve AC_2 (at productivity level ω_2) if the firm keeps the first-period input choice $\mathbf{X}_1$: Utilizing the first-period input $\mathbf{X}_1$ at the new productivity level ω_2 produces intermediary output Q_c. At both A and C, the firm does not adjust its input bundle; its *cost elasticities are the same* (points D and E). Hence, the change from A to C illustrates the change in average cost and output that is associated with change in productivity or cost shifting. On the other hand, the movement from C to B illustrates the change in output and average cost due to the change in cost elasticity (or economies of scale). On the cost curve corresponding to the new productivity level ω_2, the switch from input bundle $\mathbf{X}_1$ to $\mathbf{X}_2$ utilizes the economies of scale and leads to changes in cost and output: the firm's average cost slides from C to B, corresponding to the output expansion from Q_c to Q_2.

Regarding the specific application in the trading context, I now show that when trading happens, the cost-shifting component completely captures the cost effectiveness due to productivity heterogeneity between a quota acquirer and a quota acquiree. Let subscript "A" denote the acquirer and "E" denote the acquiree. We have:

$$\begin{aligned}\Delta Q|_{A,\Delta\omega} &\equiv \mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{At}) - \mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{A,t-1}), \\ &= \underbrace{\mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{At}) - \mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{E,t-1})}_{\text{ex-post productivity heterogeneity}} - \underbrace{\left[\mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{A,t-1}) - \mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{E,t-1})\right]}_{\text{ex-ante productivity heterogeneity}}.\end{aligned}$$

Intuitively, the gain due to within-acquirer-productivity change is the ex-post productivity heterogeneity between the acquirer and the acquiree net off the ex-ante productivity heterogeneity. The ex-ante productivity heterogeneity is the pre-trade output gap net off hypothetical deduction due to input difference:

$$\begin{aligned}\underbrace{\left[\mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{A,t-1}) - \mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{E,t-1})\right]}_{\text{ex-ante productivity heterogeneity}} &= \underbrace{\left[\mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{A,t-1}) - \mathbb{F}(\mathbf{X}_{E,t-1}, \omega_{E,t-1})\right]}_{\text{output gap between acquirer and acquiree}} \\ &\quad - \underbrace{\left[\mathbb{F}(\mathbf{X}_{A,t-1}, \omega_{E,t-1}) - \mathbb{F}(\mathbf{X}_{E,t-1}, \omega_{E,t-1})\right]}_{\text{ex-ante input gap}}.\end{aligned}$$

In the absence of trading, the "cost shifting" effect captures the cost change due to the change in technology over time within the firm.

Note that the decomposition in Proposition 6 not only highlights the separate roles of economies of scale (cost efficiency) and cost shifting (cost effectiveness), it also shows the cost-effectiveness component can be calculated without knowing information of the trade partners.

3 Background on Norwegian Cod Fishery Regulations

This paper exploits the policy transition from non-tradable cap to cap-and-trade in Norwegian cod fishery to investigate the impacts of tradability in cap-and-trade. I first estimate the production function using a variety of approaches because they rely on different identifying assumptions: OLS with fixed effects, the proxy variable approach, and the dynamic panel approach. Using these estimates, I calculate the input elasticities of output and output elasticity of cost for every vessel-year observation in the Norwegian cod fishery from 2001 to 2017. I then use difference-in-difference approach to infer the causal impacts of trading on economies of scale, in addition to other observable outcomes such as catch quantity, revenue, transacted fish sales price, production inputs. Before discussing these empirical strategies, I provide a background on fishery regulations and Norwegian cod fishery below.

3.1 An Overview of Regulations in Fisheries in General

Before going through the detail context of the Norwegian cod fishery of which data and policy experiments this paper directly analyzes, understanding the big picture of regulations in fisheries in general is useful. Similar to other common goods and public goods such as water, forests, oil, atmosphere (e.g. air pollution as a bad public good), fishery in oceans suffers the tragedy of the commons in an open-access and unregulated environment. That is, individuals in an open access system pursue their own self-interest and neglect the well-being of society, leading to overconsumption and ultimately causing depletion of the resource; see Tietenberg (2003); Costello et al. (2010); Stavins (2011); Isaksen and Richter (2019).

As a result, hundreds of fisheries have followed the lead of other natural resources and have transitioned from open access systems to property right management. The first reforming property right management is limited entry, under which only fishermen that own permits are entitled to participate in the fishery. Together with limited entry, regulators also employ other command-and-control tools—fishing season limitations, gear restrictions, area closures—to limit fishing activities. However, resource rents may be dissipated by excessive capital investment, redundant effort, or inefficient timing of harvest; see Costello et al. (2010); Stavins (2011). Note that in contrast to public goods such as air, common goods are rivaled in consumption. Fishermen may race to catch as much as possible during a limited fishing season.

Hence, regulators have stepped towards the next reform: catch share management. Under catch share management, the regulator defines the total allowable catch (TAC) of the whole fishery annually and each vessel owns a proportion of the TAC (catch share or quota) that entitles the vessel to catch up to the tonnage values of the quota. The catch share management has two types of arrangements: individual vessel quotas (IVQ) and individual transferable quotas (ITQ). In the IVQ system, a quota is attached to a boat and are not separately transferable.[2] If one wants to buy a quota, he has to buy the license and the boat. This approach is simply a command-and-control regulation in which a non-transferable cap of output or emissions is set on a facility. Although the cap can move from an owner to another owner by buying out the facility, the new cap is not allowed to combine with the existent cap to run on one facility. In contrast, the ITQ system mimics the cap-and-trade programs for air pollution. In the ITQ management, the regulator allocates shares of the harvest to individuals (individual vessels with active licenses) in the first time using a grandfathering rule and allows fishermen to trade those shares after the initial allocation.

Over the past three decades, many countries have employed catch share management. As of 2008, more than 140 fisheries in the world are managed by ITQs. Country examples include the Netherlands, Canada, Iceland, New Zealand, the U.S., Australia, Argentina, Chile, and so on; see Costello, Gaines and Lynham (2008); Chu (2009); Costello et al. (2010). For example, New Zealand was the first country to adopt ITQs as a national policy in 1986.[3] Their ITQ system is very flexible in which quotas can be divisibly traded, sold or leased, and hold in perpetuity, establishing a well-functioning market of quotas (Newell, Sanchirico and Kerr, 2005). In the U.S., most of fisheries have several different trading restrictions such as consolidation caps, sunset provisions, restrictions on leases or permanent trades, or non-use clauses for environmental participation (Grainger and Parker, 2013). On the other hand, due to quota consolidation concerns, several countries have been limiting quota trading and essentially employing IVQs such as Norway, Denmark, Sweden, the U.K., Peru; see Asche et al. (2008); OECD (2011, 2013).

Recently, Norway has gradually switched from IVQ to ITQ by allowing quota trading in certain groups of vessels in the cod fishery, the most valuable capture species in Norwegian fishery. This paper exploits this policy transition to evaluate the impacts of ITQ relative to IVQ, or cap-and-trade relative to non-tradable cap in the fishery context. I now closely discuss the regulation transition in Norway.

---

[2] Specifically, a vessel has its own license that entitles the fishery entry and the regulator defines a quota on the license. The license can be transferred and the unique quota attached with the license will thus follow the license. However, the rule is "one vessel one license (thus one quota)." Hence, "transferability" of quotas in the IVQ system is effectively transferability of licenses and boats, which is not the trading property of the cap in a cap-and-trade program.

[3] Although the Netherlands, Canada, and Iceland were the first countries to adopt ITQs in the late 1970s, New Zealand became the world's largest ITQ system in 1986 by employing the system nationally.

3.2 Norwegian Cod Fishery

The regulation transition focuses on the coastal fleet in the Norwegian cod fishery. Cod is the most valuable catch in Norwegian fishing industry. As of 2019, the primary value of cod fishing was 7.2 billion Norwegian dollars (850 million US dollars), or 34%, followed by mackerel (12%) and herring (12%).[4] Cod in Norwegian sea is Atlantic cod (*Gadus morhua*). They can live for 25 years, attain reproductive maturity between ages two and six, grow to 2m long and 40kg (88lbs).[5] Appendix A shows how they look like and the distribution along Norwegian coast.

Vessels that are allowed to fish cod in Norway are divided into two fleets: deep-sea fleet and coastal fleet. The deep-sea fleet typically consists of big commercial vessels that use active gears such as trawls and purse seines to find out the school of fish before putting the gear in the sea. The coastal fleet includes smaller vessels (less than 28 meters or have a cargo volume of less than 500m^3) that use passive gears such as yarns, long lines, hand lines, teine, net etc. that stand still in the sea and wait for the fish to reach the gear.[6] The coastal fleet has been closely monitored by the regulators, because this fleet consists of more than 2,000 vessels, contributing to main income and earnings in many communities along Norwegian coast (Nærings- og fiskeridepartementet, 2006). From 2016, the coastal fleet accounts for 97% of cod vessels and 70% of national cod quota; see Nærings- og fiskeridepartementet (2016, 2019). This paper focuses on this coastal fleet.

Table 1 summarizes the regulation changes in the cod coastal fleet. Up to 1980s, the coastal fleet in the cod fishery was able to fish freely, without restriction of access and with spacious maximum national quotas. The reason was the agreed total quota in North-East Arctic cod fishery between Norway and the Soviet Union were set significantly higher than what would be considered sustainable. After many years of over-fishing, the demand for sustainability brought up a reduction in the total quotas agreed between Norway and the Soviet Union. In 1989, the coastal fleet quickly fished up the total quota, resulting in being halted on 18 April by the Directorate of Fisheries. Such decision was a bombshell because April was in the peak of the season and many communities even had not started to go fishing. Hence, a system for individual quotas became an immediate demand and quickly supported by the Norwegian Fisher's Association.

In 1990, the cod coastal fishery moved from free fishery (with a national cap) to a closed system with individual vessel quotas (IVQ). The aim was to avoid the fishing race that had happened in 1989. Based on a minimum catch requirement in historical years, a closed group

---

[4] SSB `https://www.ssb.no/jord-skog-jakt-og-fiskeri/artikler-og-publikasjoner/fisket-verdt-21-milliardar-kroner`.

[5] See Animal Diversity Web, University of Michigan, `http://animaldiversity.org/accounts/Gadus_morhua/` and Norwegian Institute of Marine Research, `https://www.hi.no/en/hi/temasider/species/costal-cod--north-of-the-62-latitude`

[6] Before 2008, vessels in the coastal fleet had a maximum length limit of 28 meters. From 2008, the length limit in the coastal fleet was removed and replaced by the maximum cargo limit of 300m^3 (and 500^3 in 2010).

Table 1: Changes in the management of the coastal fleet in the fishery of cod in the north of 62°N

Year	Event
1980s	free fishery.
1990	limited entry (closed fishery) with IVQs.
2001	Length is recorded to legal length that is fixed, regardless of actual size upgrades.
	Fleet is divided into 4 legal length groups: 0–10.9, 11–14.9, 15–20.9, 21–27.9.
	Note that actual length has been less than 28m.
2003	Decommissioning scheme for coastal fleet up to 14.9m, from 1 Jul 2003 to 1 Jul 2009.
2004	ITQs is introduced in certain license groups of the fishery.
	Quota can be transfered between vessels in legal length group of 15–20.9m and 21–27.9m
2005	additional purchased quota has its life extended from 13/18 years to 20/25 years.
	quota ceiling increases for groups 15+, see St 2018–2019
2008	Quota trading is allowed for legal length group of 11–14.9m.
2008	change from max length of 28m to max cargo of 300m^3.
2010	change from max cargo of 300m^3 to max cargo of 500m^3.

Sources: Nærings- og fiskeridepartementet (2003, 2006, 2007, 2016, 2019); Armstrong and Clark (1997); Armstrong et al. (2014); Standal and Aarset (2008); Standal and Asche (2018).

was established. Vessels that did not satisfy the criterion could participate in an open group.[7] Vessels in the closed group were assigned individual catch shares (quotas).

Between 2001 and 2002, a length division of the coastal fleet (*Finnmarksmodellen*) was introduced to provide a fairer competition between vessels. The division categorized the coastal cod closed group into four length groups: 0–10.9m, 11–14.9m, 15–20.9m, 21–27.9m.[8] The closed group quota was divided into these four length groups before further distributed to individual vessels within a group. The intention was that vessels only competed with others within the same length group for their similar quotas. Furthermore, the physical lengths of vessels at this time were recorded into legal lengths that became fixed regardless of the size upgrades in future. The legal lengths became an attribute of the license and defined an individual quota share of a vessel in the license group.

Since 2004, quota trading has been allowed in order to reduce over-capacity and to increase profitability in the cod coastal fleet. The policy is known as a structural quota scheme (*strukturkvoteordningen*) in Norwegian regulation. The scheme allows vessels in certain license groups to trade their quotas. Due to the concern that quota trading would result in consolidation of quotas and catch into a few hands of fishermen, the trading scheme were implemented in vessels in only the two upper groups, 15–20.9 and 21–27.9, beginning on 1 January 2004. In 2007, after a review of the legislation, the scheme was expanded to cover vessels in the license group of 11–14.9, starting from 1 January 2008.

The scheme also imposes several restrictions on quota trading. First, vessels are only al-

---

[7] The open group is regulated by a total group quota rather than individual quotas and this group quota is substantially low, only about 5% to 10% of the cod fishery quota; see Nærings- og fiskeridepartementet (2019).

[8] Until 2008, vessels in the coastal fleet must be shorter than 28 meters.

lowed to trade quotas within the same license group. Second, the vessel that sells the quota must exit the fisheries permanently (by being scrapped or sold). Third, a portion of the transferred quota (20%) must be deducted and given to the other vessels in the group. Fourth, the transferred quota is only valid for limited time, 13 years if being sold (15 years if the vessel is scrapped).[9] The final rule is a geographical restriction: Vessels in the south are not allowed to buy quotas from the north, although the North can buy quotas from the South.

Overall, the quota trading scheme in Norwegian cod coastal fishery has switched IVQs to ITQs in certain licensed length groups, 15–20.9 and 21–27.9 (and 11–14.9m), since 2004 (and 2008). As discussed in a pubic report by the Ministry of Trade and Industry (Nærings- og fiskeridepartementet, 2006), the goal of the transition is to reduce overcapacity and increase profitability in the cod fishery. They justify that IVQs were able to prevent the expansion of further overcapacity, but IVQs would not give the industry incentives to remove existing overcapacity. Under an ITQ system, fishermen with a quota would like to have quasi-property rights to a certain proportion of the total quota. If overcapacity led to reduced profitability in the industry, quota could be bought and sold until costly overcapacity is gone. The least efficient vessels will be taken out of fishing in exchange for the more efficient ones increasing their catch. This vision of reducing overcapacity of the regulators explains the motivation for the scrapping condition that requires vessel that sells the quota give up its whole quota and exit the cod fishery permanently.[10]

This paper exploits the policy transition from IVQs to ITQs in only certain groups in the coastal cod fleet to study the effects of ITQs, relative to IVQs, on the vessel-level performance in the fishing market. The comparison between ITQs and IVQs is analogous to comparing cap-and-trade to command-and-control. In theory, cap-and-trade performs better to achieve economic efficiency goal by minimizing costs of compliance in the case of emissions reduction, or by maximizing the value of resources in the case of natural resources. In fishery, as revealed by the above stated vision of the Norwegian Ministry of Trade and Industry, ITQs are expected to increase efficiency and profitability. I will test this hypothesis by exploring the impacts of ITQs on productivity, fish sale prices, and production costs. I will also test whether the quota trading results in increase in market power due to consolidation concern.

It is worth mentioning that the committee in 2003 proposed two measures to reduce the number of vessels in the coastal fleet: the ITQs system and the decommissioning scheme. Whereas the ITQs were implemented for vessels with legal lengths from 15m, the vessels below 15m (in terms of legal length) were subject to the decommissioning scheme in which regulator paid out the fisherman to buy back the vessel and license. The decommissioning scheme carried out from 1 July 2003 to 1 July 2009. This scheme led to a substantial reduction

---

[9] This valid duration was implemented for the scheme from 2004. After a legislation review in 2007, the valid durations became 20 years if being sold and 25 years if the vessel is scrapped.

[10] The scrapping condition may imply that quotas are not traded divisibly. However, given that the trade happens between vessels within the same license group and vessels in a license group may have different quota due to different legal lengths, the whole policy scheme has certain degree of divisibility.

in the number of vessels in the control group, license of 0–14.9, which would raise a caveat for the interpretation of the empirical results using a difference-in-difference approach. The next section discusses the empirical approach and identification in detail.

4 Empirical Strategies

The goal of this paper is to examine the effect of cap-and-trade program in Norwegian fishery on individual vessel performance in the product market, e.g. fish harvest and sales. To do that, I use difference-in-difference approach to compare the change in a variety of fishing outcomes between the treatment group and control group and between before and after the policy. I first look at the policy impact on standard fishing outcomes such as harvest quantity and revenue. I then investigate the impact on other observable outcomes: transacted fish sales prices, vessel length, crew size, distance from fishermen's home municipality to major catch location, and the number of fishing trips in a year. Finally, I look at the impacts on productivity and economies of scale. Section 4.1 discusses the difference-in-difference strategy that estimates the causal impacts of trading on these outcomes.

Because productivity and economies of scale are not observed, I estimate them by estimating a production function and using Proposition 2. Section 4.2 discusses the approaches estimating a production function. Section 4.3 discusses the identification of all empirical strategies.

4.1 Estimating the Causal Effects of Cap-and-trade

I estimate the effect of cap-and-trade relative to non-tradable cap by exploiting the fact that only certain vessels are allowed to trade quotas. Vessels that hold licensed lengths below 11 meters are never allowed to trade and hence considered in a control group. The other vessels are in treatment groups with staggered adoption.[11] Licensed length group 2 (11–14.9m) can trade quotas from 2007 and licensed length groups 3 and 4 (15–20.9m and 21–27.9m) can trade quotas from 2004. My identification strategy is to estimate difference-in-difference specifications comparing the difference between treatment and control groups and the pre-treatment and treatment-periods. Note the choice of buying quotas is voluntary. Because not every vessels in the trade-qualified group chooses to buy quotas, I estimate the intent-to-treat effect (ITT) using a difference-in-difference (DID) specification and the local average treatment effect (LATE) using a DID instrumental variable specification.

---

[11] The staggered DID estimate may be biased because treated individuals get treated at different time, making their presence in the sample unbalanced. Hence, one would have to correct the estimate; see Callaway and Sant'Anna (2020); Goodman-Bacon (2021). In this paper, for the intent-to-treat of the trading program, we only have three treatment groups. So, we can illustrate the effects of each group separately relative to a clearly defined control group and time.

Intent-to-treat

I estimate the following DID specification to identify ITT:

$$Y_{it} = \beta_{ITT} \textit{Trade Qualified}_{it} + \eta_i + \tau_t + \epsilon_{it}, \tag{23}$$

where Y_{it} is an outcome of interest, such as logged revenue, logged catch quantity, sale prices, fishing frequency, productivity, economy of scale, and production factors of vessel i at time t. The variable *Trade Qualified*$_{it}$ equals 1 if vessel i is in the trade-qualified group at time t, and zero otherwise. Equation (23) includes vessel fixed effects η_i to account for permanent differences in the operating skills of time-invariant ownership during the period 2001–2017. The model also includes time fixed effects τ_t to adjust for the average effects of time-varying factors (e.g. weather, sea temperature, stock levels, seasonality) that generate variation in the outcome of interest across all vessels.

The parameter of interest is β_{ITT}. It is a DID estimator that compares the change in outcome Y of trade-qualified vessels after trade qualified status to before, relative to the vessels that are not qualified for quota trading (licensed lengths below 11 meters). This is the intent-to-treat effect that measures the impact of the trade-qualification program.

The identification assumption of this DID approach is that, conditional on the fixed effects, differences between trade-qualified vessels and non-qualified vessels are on average similar for pre-trade-legalization and post-trade periods had the vessels not traded quotas. This assumption is untestable. However, the parallel trend assumption in the counterfactual condition can be plausibly to believe if the two groups have parallel trend in the pre-trade-qualification period. I, hence, use the event study with lags DID coefficients to test whether there is no difference in the outcome of interest between the two groups in pre-trade-qualification period, relative to the year the program is implemented. This event study also helps inform any anticipatory effects of the trade program. If fishermen expected the trade program would be enacted, they would adjust their production factors before the program officially enacted, showing a deviation in fishing outcomes from the parallel trend for years just right before the policy enforcement. Besides the lags DID coefficients, I also include the leads coefficients in the event study to explore the dynamic effects of the policy. Results in sections 6 and 7 confirm the parallel trend in the pre-trade period and shows the effects of trading on catch quantity and revenue happened immediately after trading was allowed. Fishermen immediately utilize additional quotas by going fishing more often whereas take time to invest in capital (build bigger vessels).

Despite the pre-trade-period parallel trend verification, a threat to the identification of the causal effect is the spillovers of the trading impacts on vessels in the control group. Because the trading requires one of the two vessels in the transaction leave the fishery, vessels in the control group may benefit by facing fewer competitors both in the fishing ground and in the fish sales market. Specifically, vessels in the control group may catch more given lower congestion costs. Figure 4 in the next section shows the license group 0–10.9m, as a control group, goes fishing within 150km from the coast, which are distinctly distant from the fishing locations of vessels in the other groups, thereby alleviating the spillovers due to congestion costs. However, a small threat of spillovers in harvest may happen if fishing activities of big vessels in faraway locations in the ocean interfere the migration of cod. In that case, vessels in the control group may benefit from the fact that there would be fewer vessels, making the DID estimator underestimate the policy impact on trade-qualified vessels' harvest. Similarly, in the landing market, fewer competitors may help vessels in the control group sell their fish at higher prices than before, thereby causing the DID estimator underestimate the market-power-abuse impact of consolidation on fish prices.

Average treatment for the treated

Because the trade-qualification program is not mandatory for all vessels in the trade-qualified group, the DID approach in equation (23) estimates the effect of the qualification program rather than the treatment effect of acquiring quotas through the program. I now use a DID instrumental variables (IV) specification to identify a Local Average Treatment Effect (LATE) of acquiring quotas. Specifically, I estimate the following equation:

$$Y_{it} = \beta_{LATE} \mathit{Quota\ Acquisition}_{it} + \eta_i + \tau_t + \epsilon_{it}, \tag{24}$$

where $\mathit{Quota\ Acquisition}_{it}$ equals one for all periods after vessel i buys all quotas from another vessel in the trade-qualified group and zero otherwise. This indicator is instrumented by the trade-qualification program status $\mathit{Trade\ Qualified}_{it}$ to reduce selection bias.

The coefficient β_{LATE} measures the Local Average Treatment Effect (LATE). It measures the average changes in Y_{it} from quota acquisition on "complier" vessels that will acquire quotas whenever they are qualified for trading. In this context with one-sided noncompliance, we only have either never-takers or compliers. Hence, the LATE would be also the average treatment effect on the treated. The key identification assumption to interpret β_{LATE} as a causal effect is the exclusion restriction that requires the trade-qualification affect the outcome of interest only indirectly via an effect on trading execution. However, the Norwegian quota trading program by design sets a rule in which the quota acquirer may only take 80% of the quota and leave 20% equally shared to the other vessels in the group. Hence, there are vessels that would never trade but be able to gain higher quotas due to the program, causing a potential violation of the exclusion restriction condition. In this case, the simpler DID estimation of 24 using OLS with fixed effects may give more convincible estimates of the quota acquisition impacts. The causal interpretation of this (OLS) average treatment for the treated requires the vessels that acquire quotas have similar trends in fishing performances to other vessels had the acquisition not happened. While this assumption cannot be completely tested, as in the case of intent-to-treat, I provide an event study that carefully looks at the changes in the vessel performance right before and after the quota acquisition to test the parallel trend patterns in the pre-acquisition period.

4.2 Estimating a Production Function and Productivity

I now discuss the estimation of a production function to obtain productivity and output elasticity of total costs. These two estimated variables become additional dependent variables in the above difference-in-difference framework to analyze the impact of trading policy on productivity and economies of scale. However, the variation in productivity induced by the trading policy must be separated from the variation in inputs in the production function estimation process, before being fed in the difference-in-difference regressions.

In this paper, I consider the yearly production function of a vessel. Production factors include vessel size (length) K_{it}, crew size (labor) L_{it}, distance from the fishermen's municipality to major catch location D_{it}, and the number of trips in a year M_{it}. The reason is that the exact quota tonnage is set annually. Every year, the regulator decides the total allowable catch for the whole fishery and defines the conversion factor that converts a vessel's quota share to his quota tonnage. Quotas are not bankable. Hence, it is reasonable to assume the vessel's owner decides input factors that are critical for a production year instead of a trip.

Consider a general production technology with four factors and Hicks-neutral productivity:

$$Q_{it} = F(K_{it}, L_{it}, D_{it}, M_{it}; \beta) \exp(\omega_{it}), \tag{25}$$

where $\exp(\omega_{it})$ is the productivity of vessel i in year t (logged quantity-based total factor productivity or logged TFPQ). In an empirical framework, we observe logged output y_{it} and assume $y_{it} = \ln Q_{it} + \epsilon_{it}$, where ϵ_{it} is an exogenous unexpected shock to production. I estimate the following equation:

$$y_{it} = f(k_{it}, l_{it}, d_{it}, m_{it}; \beta) + \omega_{it} + \epsilon_{it}, \tag{26}$$

where $k_{it}, l_{it}, d_{it}, m_{it}, \omega_{it}$ are logged inputs and logged neutral productivity.

Because the logged productivity ω_{it} is unobserved to an econometrician, there are two challenges in estimation. First, a vessel's owner choose their inputs based on the realization of ω_{it}, causing simultaneity bias. Second, vessels that exit over time are those that have low productivity, causing selection bias. To address these challenges, literature has suggested four solutions: using input prices as instrument variables, using OLS with fixed effects, using the proxy variable (control function) approaches (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg, Caves and Frazer, 2015; Gandhi, Navarro and Rivers, 2020), and using dynamic panel approaches (Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998, 2000). In this paper, I use OLS with fixed effects, a proxy variable approach, and a dynamic panel approach.

4.2.1 Proxy Variable Approach

The proxy variable approach relies on a set of assumptions on timing decisions of inputs, the relation between a proxy variable and the scalar unobservable (productivity), and the evolution of productivity over time. Typically, the proxy variable approach in the literature assumes capital is predetermined, that is, capital in today's period t is determined in the previous period $t-1$ through the past adjustment in investment. In this paper, given the Assumption 1, I assume all inputs are variable. As you will see, this affects the choice of instruments and moment equations I use in the estimation. I now firstly describe the estimation method in this paper and discuss the relation with the literature in Section 4.3 (Identification).

Together with Assumption 1, the following Assumption 2 is needed for the estimation of a production function using the proxy variable approach in this paper.

Assumption 2. The proxy variable approach to estimate a production function in this paper assumes Assumption 1 and

(2.1) The conditional demand for a proxy variable input, $m_{it} = \mathcal{M}_{it}(k_{it}, l_{it}, d_{it}, \omega_{it})$, is strictly monotone in a single unobservable ω_{it}.

(2.2) The productivity ω_{it} evolves in a Markovian process:

$$\omega_{it} = g(\omega_{i,t-1}, \textit{Trade Qualified}_{i,t-1}) + \xi_{it},$$

where ξ_{it} is an exogenous shock in productivity that is uncorrelated with information at $t-1$, that is $\boldsymbol{E}[\xi_{it}|\mathbb{I}_{t-1}] = 0$.

Under the Assumption (2.1), the conditional demand for a proxy input is inverted and substituted into the production function to get

$$y_{it} = \underbrace{f(k_{it}, l_{it}, d_{it}, m_{it}; \beta) + \overbrace{\mathcal{M}_{it}^{-1}(k_{it}, l_{it}, d_{it}, m_{it})}^{\omega_{it}}}_{\equiv \chi_{it}} + \epsilon_{it}. \tag{27}$$

Hence, in the first stage, I run the following regression

$$y_{it} = \chi_{it}(k_{it}, l_{it}, d_{it}, m_{it}) + \epsilon_{it}, \tag{28}$$

to obtain estimates of expected output $\widehat{\chi}_{it}$. Although the coefficients in the first stage are not the coefficients for the production function, the goal is to separate productivity from shock ϵ_{it}.

In the second stage, I estimate the production function coefficients β using Assumption (2.2). With this assumption, I can compute productivity for any value of β, using $\omega_{it}(\beta) = \widehat{\chi}_{it} - f(k_{it}, l_{it}, m_{it}; \beta)$. By nonparametrically regressing $\omega_{it}(\beta)$ on its lag and event variables affecting productivity, I recover the exogenous shock $\xi_{it}(\beta)$. With the timing of the firm's decisions on k, l, d, m and the uncorrelation between exogenous shock ξ_{it} and past information, I can use the following moments to estimate β:

$$\mathbf{E}[\xi_{it}(\beta)x_{i,t-1}] = 0, \tag{29}$$

where x is the input vector (k, l, d, m).

For the specification of the production function, specifying a flexible enough production function is important to obtain a individual-specific and time-varying input elasticities of output. In this paper, I estimate the translog production function. Given GMM estimates of β in the second stage, the output elasticity for capital, for example, is given by

$$\widehat{\theta}^K_{it} = \hat{\beta}_k + 2\hat{\beta}_{kk}k_{it} + \hat{\beta}_{kl}l_{it} + \hat{\beta}_{kd}d_{it} + \hat{\beta}_{km}m_{it}. \tag{30}$$

After getting all input elasticities of output, I can calculate output elasticity of total cost $\widehat{\phi}_{it} = (\sum_{X\in\mathbb{X}}\widehat{\theta}^X_{it})^{-1}$.

4.2.2 Dynamic Panel Approach

Compared to the proxy variable approach, the dynamic panel approach is more flexible because it does not require Assumption (2.1) (scalar unobservable of ω and strict monotonicity of the proxy input demand) and it can allow additional unobservable fixed effects that affect the production function. However, it requires the serial correlation in ω_{it} is linear.

Assumption 3. The dynamic panel approach to estimate a production function in this paper assumes the productivity evolves as follows: $\omega_{it} = \rho\omega_{i,t-1} + \alpha \textit{Trade Qualified}_{i,t-1} + \xi_{it}$, and considers broadly the production function $y_{it} = f(k_{it}, l_{it}, d_{it}, m_{it}; \beta) + \omega_{it} + \eta_i + \delta_t$.

With this assumption, we have

$$y_{it} = \rho y_{i,t-1} + f_{it} - \rho f_{i,t-1} + \alpha \textit{Trade Qualified}_{i,t-1} + \delta_t - \rho\delta_{t-1} + \underbrace{(1-\rho)\eta_i}_{\eta_i^*} + \underbrace{\xi_{it} + \epsilon_{it} - \rho\epsilon_{i,t-1}}_{\epsilon_{it}^*},$$

where $f_{it} = f(k_{it}, l_{it}, d_{it}, m_{it}; \beta)$. According to Arellano and Bond (1991), I can first difference the above equation to eliminate the individual fixed effects, obtain the residual $\Delta\epsilon^*_{it}$, and proceed the estimation using the moment condition $E[Z_{i,t-s}\Delta\epsilon^*_{it}] = 0$, where the instruments $Z_{i,t-s}$ are the s lagged values of input factors k, l, d, m and output y. The exogeneity of these instruments is owing to the orthogonality between exogenous shock ξ_{it} and past information and between current shock ϵ_{it} and current input choices. A concern about these instruments is that they may be weakly correlated with the first differenced residuals, raising poor performance (bias and imprecision) in a finite sample; see Blundell and Bond (1998). Hence, I follow the suggestions by Blundell and Bond (1998, 2000) to include additional moment condition $E[\Delta Z_{i,t-s}(\eta^*_i + \epsilon^*_{it})] = 0$. This additional moment requires further assumptions that $E[\Delta x_{it}\eta^*_i] = 0, E[\Delta y_{i2}\eta^*_i] = 0$.

4.3 Identification

Identifying the causal impacts of the trading policy on productivity and production cost relies on two main identifying strategies. The first is to identify productivity and output elasticity of total cost using either the proxy variable approach or the dynamic panel approach. The second is to identify the causal impacts of trading using the difference-in-difference identifying conditions.

4.3.1 Identifying the Production Function and Cost Elasticity

Under the proxy variable approach in this paper, the gross output production function and cost elasticity are identified upon the Assumptions 1 and 2. Compared to the literature, there are two noticeable differences. First, I assume capital is variable, whereas the literature has assumed capital is predetermined (and dynamic); see Olley and Pakes (1996); Levinsohn and Petrin (2003); Ackerberg, Caves and Frazer (2015); Gandhi, Navarro and Rivers (2020). With this difference, instead of using the moment condition $E[k_{it}\xi_{it} = 0]$, I use the condition $E[k_{i,t-1}\xi_{it} = 0]$. As discussed in Ackerberg, Caves and Frazer (2015), if capital k_{it} is predetermined and depends on the investment adjustment in the previous period $t-1$, the current capital will be uncorrelated with exogenous shock ξ_{it}. For a variable input x_{it} that is chosen today and correlated with today's productivity, the past value of the variable input can be used instead in the moment condition. The coefficients of input variables are not identified in the first stage, but they are identified in the second stage of GMM upon the uncorrelation between the exogenous shock of innovation ξ_{it} and past information.

Second, I estimate the gross output production function, whereas Ackerberg, Caves and Frazer (2015) estimate the value added production function.[12] Whereas Ackerberg, Caves and Frazer (2015) exclude the proxy input m_{it} in the production function, including it in the production function is important in my case because I need to obtain all elasticities of inputs with respect to output. The exclusion of the proxy input in the production function was due to the concern about the nonidentification of the gross output production function using the proxy variable approach as being raised by Gandhi, Navarro and Rivers (2020). However, Gandhi, Navarro and Rivers (2020) formally show the nonidentification of the gross output production function in the absence of time-series variation in relative prices (of input and/or output). As a result, the cross-sectional variation in the proxy variable is not enough to help identify the gross output production function. As shown in the below descriptive statistics (Figure 4), this is not the case in this paper: I consider the number of fishing trips in a year of a vessel as the proxy variable and this variable has variation both across vessels and time (year).

Ultimately, the key assumption that leads to the nonidentification of the gross output pro-

---

[12] Olley and Pakes (1996) estimate the value added production function. Levinsohn and Petrin (2003) estimate the gross output production function but they estimate the coefficient of labor in the first stage, which is shown to be biased in Ackerberg, Caves and Frazer (2015).

duction function shown by Gandhi, Navarro and Rivers (2020) is the Assumption (2.1) that assumes productivity is the only scalar unobservable and the conditional demand for the proxy input is strictly monotone in productivity. The dynamic panel approach can avoid this assumption and can be used to estimate a gross output production function. This approach also allows other unobservables in the forms of fixed effects, besides productivity, affect the production function. Hence, I also consider the results using this approach. The limitation of this approach is to assume the linearity of the serial correlation in productivity (see Assumption 3).

Section 6 reports the main results using the proxy variable approach to estimate the production function. Appendices report results using OLS with fixed effects and the proxy variable approach for the production function estimation. Results show the cost elasticity and productivity by the three approaches exhibit some differences in magnitudes but follow similar distribution shapes by vessel groups and years.

4.3.2 Identifying the Causal Impacts of Trading

The causal impacts of trading are identified upon the principle of difference-in-difference: assuming that the group of trade qualified vessels and the group of non-qualified vessels follow similar (parallel) trends in productivity and economies of scale, and that the impacts of trading do not spill over to non-qualified group, as noted in Section 4.1.

When combining with the estimation of a production function, separating the impact of trading on productivity from the impact on input factors *in the production function estimation step*, before feeding into the difference-in-difference step, is important. Otherwise, the impact of the policy (trading) on productivity would be confounded with the impact of policy on input choices (and cost elasticity). To do such separation, including both $\textit{Trade Qualified}_{i,t-1}$ and $\omega_{i,t-1}$ in the evolution of productivity is the key and solves two issues.[13] First, ignoring $\textit{Trade Qualified}_{i,t-1}$ in the evolution process would let ξ_{it} absorb the trading impact. If trading also affects input choices $(k_{i,t-1}, l_{i,t-1}, d_{i,t-1}, m_{i,t-1})$, then ξ_{it} would correlate with the input choices. Second, more-productive firms tends to self select to the trading program. Including lagged productivity $\omega_{i,t-1}$ (together with $\textit{Trade Qualified}_{i,t-1}$) helps control the potential self-selection of trading.

In summary, the estimated productivity in the estimation stage of a production function includes its own impact of trading without being confounded by the impacts of trading on produc-

---

[13] This productivity process has been noted by De Loecker (2013); Braguinsky et al. (2015) to study the effects of a firm's export status and plant acquisition on productivity, respectively. The original proxy variable approach by Olley and Pakes (1996); Levinsohn and Petrin (2003); Ackerberg, Caves and Frazer (2015) assumes the Markov productivity evolution without controls: $\omega_{it} = g_1(\omega_{i,t-1}) + \xi_{it}$. De Loecker (2013) also notes that such a productivity process as in (27) can directly estimate the impact of the event, trading in my context. A flexible specification of g can also offer the distribution of the heterogeneous impact of trading. However, I, similar to Braguinsky et al. (2015), take a two-stage approach where the second stage is the difference-in-difference regression. Although the two-stage approach delivers the average effect of trading instead of vessel-specific effect, this approach offers two advantages. First, I want to look at the changes from before to after the trading event not just for the trade-qualified vessels only, but also in comparison to a control group. Second, I want to use a consistent framework to investigate multiple outcomes.

tion inputs. The estimated production coefficients β also keep their own impact of trading and, together with the variation in production inputs, contain the impact of trading on economies of scale (cost elasticity). In other words, whereas variation in each production input identifies the impact of trading on the input itself (if any), the total variation in *all* inputs weighted by the production coefficient (after being separated from the variation in productivity) identifies the impact of trading on economies of scale.

5 Data

Table 2: Summary statistics

	count	mean	sd	min	max
Panel A: Sample of trip-level observations					
catch quantity (tonne)	1,158,557	1.61	4.04	0.00	224.62
revenue (thousand NOK)	1,158,557	24.24	65.24	0.00	4345.58
price (NOK/kg)	1,158,557	15.79	6.47	0.20	4955.00
crew (person)	1,158,557	2.11	1.49	1	99
distance (km)	1,158,557	133.78	261.43	0.05	2318.62
Panel B: Sample of yearly observations					
catch quantity (tonne)	30,776	60.83	94.98	0.00	1874.40
revenue (thousand NOK)	30,776	915.26	1424.22	0.01	28250.42
average value (NOK/kg)	30,776	16.10	5.66	4.79	128.83
length (m)	30,776	12.86	4.93	4.25	55
crew (person)	30,776	2.24	1.62	1	21.25
distance (km)	30,776	180.01	286.38	0.05	1642.74
# trips	30,776	37.68	23.80	1	213

The data are given by the Norwegian Directorate of Fisheries under a data confidentiality agreement. The data consist of four sets for the Norwegian cod fishery from January 2001 to December 2017. First, the vessel registry records the yearly registration status of a vessel and its physical characteristics including length, engine power, tonnage, and built year. Second, the ownership registry describes the identity of a vessel's owner and their name, address, and organization type as of 31 December every year. Missing ownership is filled in using complement files of vessel events on changes in owner and vessel identification for continuous years. Third, the license registry records the license information and its valid duration a vessel holds. Fourth, the landing data record transactions of first hand sales of fish between a fisherman and a buyer (typically processing firms). The recorded transaction includes information on catch quantity, unit price, fishing vessel, landing date, the latest catch date, major catch location, fishing gear, crew size, and landing municipality. Unfortunately, trip duration is not available. When addressing fishing frequencies, I suppose each latest catch date represents for a trip in the corresponding week and month.

I merge four datasets to compile two main samples for the analysis: trip-level sample and yearly sample. Table 2 shows the summary statistics of variables in the two samples.[14] Panel A summarizes the trip-level sample. Key variables include catch quantity, transacted unit price, revenue (as a product of quantity and unit price), crew size, and distance from the major catch location to the home municipality of the fishermen (also the vessel's register). These variables are recorded for every landing transaction. On average, a vessel catches 1.6 tonne (1.7 US ton), but there is a big gap across vessels. Some vessels catch only several kilograms of cod and some catch up to 224 tonnes (247 tons). The unit price has the mean value of 15.79 NOK/kg (82 cents/lb) and also varies a lot by transactions. Crew size on average is 2. There are 14 observations of 99 people on board, whereas the second highest value of the crew size is 61. Although 99 is definitely not the code of missing values, it may be a misreport by the fishermen. Because there are only such 14 observations and it is important to keep the records of catch quantity, I decide to keep these observations in the sample for analysis.

Panel B summarizes the yearly observations. The yearly catch quantity and revenue are the sums of trip-level values for each vessel. The crew size and distance are the averages of trip-level values, weighted by the trip quantity.

For the main analysis, I use the yearly sample, except the investigation of transacted prices. Main fishing outcomes are catch quantity and revenue. Production inputs include vessel length, crew size, fishing distance (from fishermen's municipality to major catch location), and the number of trips in a year.

Figure 4 plots the average values of key variables, namely the number of vessels, catch quantity, revenue, vessel length, crew size, fishing distance, and the number of trips by licensed length group over years. Panel 4a shows the number of vessels over years. In all license groups, the numbers of vessels decrease over time. For vessels with licensed lengths from and above 15 meters, the number of vessels in these two groups has decreased since 2004, after the trading program started and allowed a vessel to buy out quota of another vessel that had to exit after selling the quota. Similarly, the number of vessel with licensed length of 11–14.9m has decreased after 2008. Between 2003 and 2008, a condemnation program in which the state bought back a number of vessels with licensed lengths below 15m played a key role to dramatically reduce the number of vessel in this group. The fall in the number of vessel in the license length 0–10.9m after 2010 is attributed to voluntary exits rather than the condemnation scheme or the trading policy. Note that although the number of vessel in the non-tradable group (0–10.9m) reduces significantly over years, the percent number in fact increases, see panel 4b Hence, the falls in the numbers of vessels in the tradable groups after the trading policy enacted are more substantial relative to the non-tradable group.

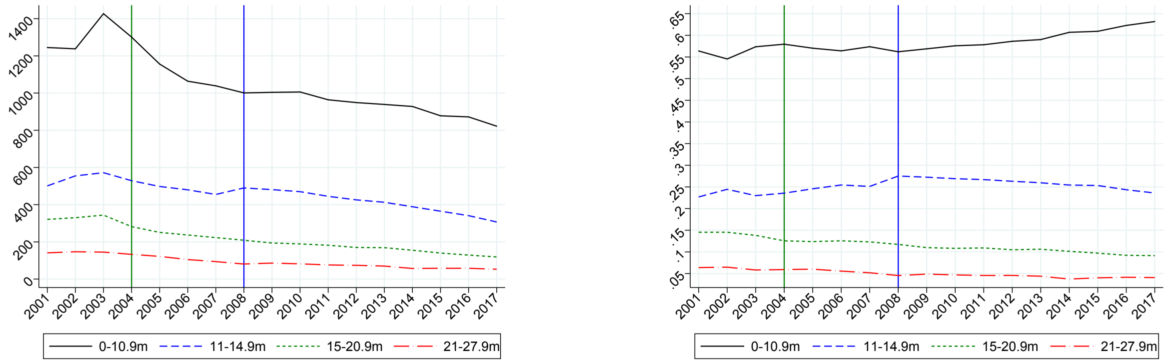
Panels 4c and 4d show vessels in a bigger licensed length group catch and earn more systematically. The tradable license groups catch significantly higher after the trading program applied.

---

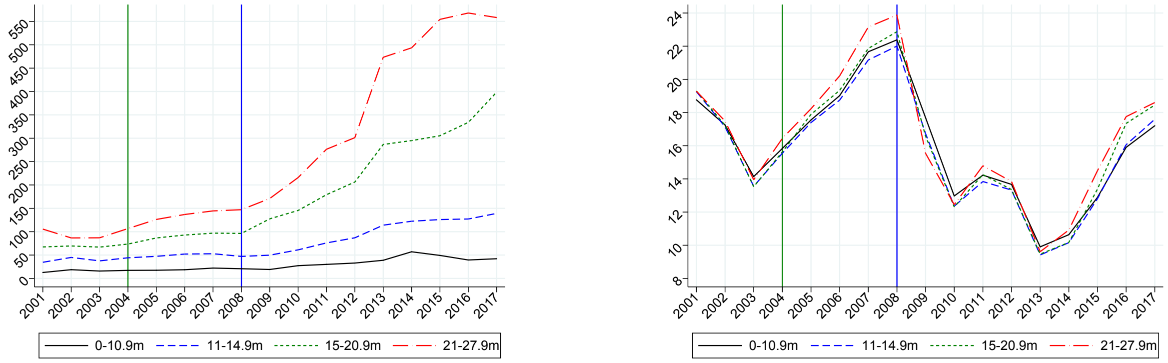
[14] One useful note is 10 nok $\approx$ 1 euro $\approx$ 1.15 usd. 1 kg $\approx$ 2.2 lbs. 1 tonne $\approx$ 1.1 US tons.

Panels 4e–4h show the trend in yearly production factors: vessel actual length, crew size, distance from fishermen's municipality to major catch location, and the number of trips. Vessels in a bigger licensed length group systematically have certain longer vessel and bigger crew size. Since 2010, vessels have been allowed to be longer than 28 meters as long as their cargo sizes are below 500m^3. We see vessels in the license group 21–27.9m have taken this advantage and expanded their sizes beyond 28m. Given the big gap in actual vessel size among license groups, one may concern vessels in the big license group would crowd out small vessels in the fishing ground. Panel 4g alleviates this concern by showing that vessels in different license groups fish in areas distant from each other. Although the two license groups, 11–14.9m and 15–20.9m, fish in overlapped areas, the group 0–10.9m as a control group fishes within 150km from the vessel home municipality's coast, distinctly not overlapping with other groups. The separation in fishing ground between tradable group and control group is important, because it helps mitigate the spillovers that threat the identification of the difference-in-difference estimation design in my context. A small threat of spillovers in harvest may still exist if the fishing activities of big vessels in faraway locations interfere the migration of the cod school. Figure A2 in Appendix A shows the distribution area and spawning area of cod in the north of the latitude $62°N$, the whole fishery that is studied in this paper. The spawning areas are near to fjords and the coastal areas, where small vessels fish. Big vessels fishing in the ocean may interfere with the migration, but there exists several management measures (closure areas and timing restrictions) during a year to limit fishing activities during the spawning period.

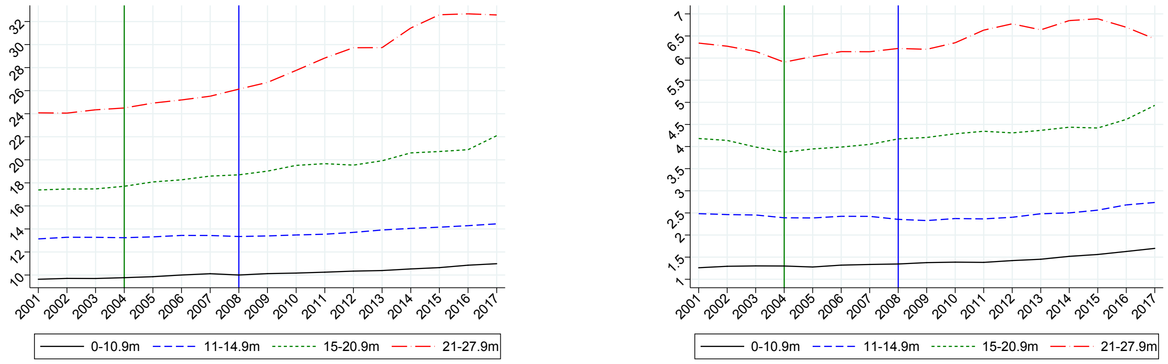
Figure 4h shows vessels in the big license groups go fishing less often than the ones in the smaller license groups. There may be two reasons. One is vessels in the big license groups travel on much bigger vessels. Another reason is the figure shows the number of trip in a year. Big vessels may prolong the trip thanks to higher safety and better equipped. Unfortunately, the trip duration is not reported. Hence, instead of investigating the trip production, I focus on the yearly production with four yearly inputs: length, crew size, fishing distance, and the number of trips. In fact, focusing on yearly production is also plausible for the reason that these four inputs are potentially all main factors a fisherman (or vessel's owner) would consider to design a plan to utilize the *yearly* quota.


(a) # vessels

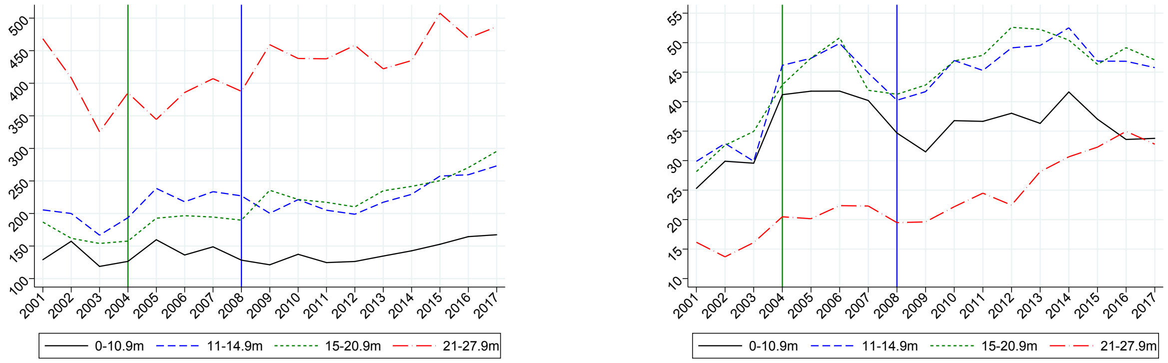
(b) # vessels in percentage


(c) catch quantity (tonne)

(d) price (NOK/kg)


(e) length (m)

(f) crew size


(g) fishing distance

(h) # trips

Figure 4: Description of key variables by licensed length group in a yearly-observation sample

6 Results

6.1 Catch Quantity and Revenues

Table 3 presents the effects of CAT program and quota acquisition on the yearly fishing performances: catch quantity and revenue. All specifications include year fixed effects, vessel fixed effects, and owner fixed effects. Panel A shows the results for the difference-in-difference (DID) specification of equation (23) that estimates an intent to treat (ITT) effect. Across the columns, the estimates suggest that the cap-and-trade program, compared to the previous non-tradable cap, increases the harvest by 8.5% and revenue by 8%. Panel B separates the ITT by tradable licensed length group. Note that vessels are divided into four groups depending on licensed lengths: below 11m, 11–14.9m, 15–20.9m, and 21–27.9m. The latter two groups, 15–20.9m and 21–27.9m, are allowed to trade quotas from 2004. The licensed length group 11–14.9m may trade quotas from 2008. Results show that the three biggest groups improve harvest and revenue by roughly 23%, 13%, and 4%, respectively.

Panels C–F estimate the average treatment on the treated (ATT) of quota acquisition using DID specifications of equation (24). Panels C and D show the results of the OLS specifications with fixed effects (OLS ATT). Panels E and F report the local average treatment effect (LATE) using instrumental variable specifications with fixed effects. This is the average treatment effect of the trading program on compliers. As expected, LATEs are greater than ITTs. LATEs are also greater than the ATT in this case, because not all vessels in the trade qualified groups are compliers. Results show LATEs estimate compliant vessels that acquire quotas catch 40% more than before and earn 37% higher. The ones in the biggest licensed length group, thanks to higher quota ceilings, even double their revenues. The ones in the small licensed length group (11–14.9m) earns 26% additional revenues.

Figures 5 presents event-study graphs for the policy effects on catch quantity and revenue. Panels A and B show the relative changes between vessels in the trade-qualified groups and vessels in the never-treated group (licensed length below 11m) before and after 2004. Year 2004 is the first time when the trading program was introduced. We see that vessels in the trade-qualified groups improve their harvest and revenues modestly between 2004 and 2008. Since 2008, the trade-qualified groups dramatically increase their harvest and earnings. Since 2013, some vessels in the licensed length 21–27.9m even double their harvest and income.

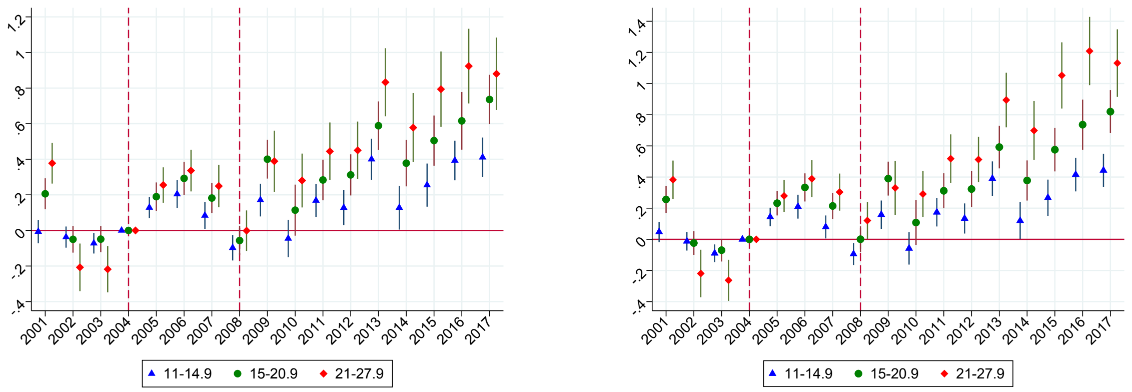
Panels C–F illustrate the event studies for the effect of quota acquisition. They plot the coefficients of the regression (24), except the indicator $\textit{Quota Acquisition}_{it}$ is interacted with dummies for years before and after the acquisition. The one year prior to the quota-acquisition year is normalized. Panels E and F show the effect of quota acquisition by licensed length group. The figures show coefficients of interactions for years before acquisition are very close to zero, implying parallel trend assumptions satisfied. This suggests the trends in catch quantity and revenue of vessels that acquired quotas are very similar to the trends of vessels that did not.

Table 3: Effects of trading policy on catch quantity and revenue

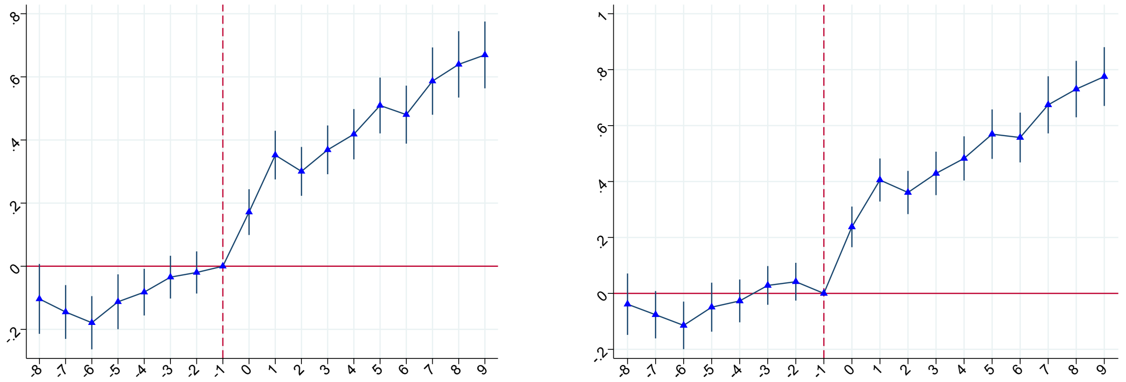
	(1) weight (tonne)	(2) logged weight	(3) revenue (tonne)	(4) logged revenue
Panel A: ITT (pooling all trade qualified groups)				
Trade qualified	9.579***	0.085***	131.271***	0.080***
	(1.611)	(0.027)	(22.992)	(0.027)
Panel B: ITT by trade qualified group (license group)				
21-27.9m × From 2004	39.611***	0.226***	671.789***	0.247***
	(6.908)	(0.047)	(114.287)	(0.050)
15-20.9m × From 2004	19.357***	0.127***	310.356***	0.128***
	(3.470)	(0.033)	(62.636)	(0.039)
11-14.9m × From 2008	0.580	0.044	-31.895	0.032
	(2.606)	(0.035)	(28.960)	(0.034)
Panel C: pooled ATT using DID FE				
Quota acquisition	72.483***	0.464***	917.469***	0.450***
	(4.594)	(0.023)	(64.809)	(0.021)
Panel D: ATT by license group, using DID FE				
Quota acquisition × 21-27.9m	157.203***	0.623***	2217.572***	0.612***
	(19.930)	(0.056)	(287.735)	(0.056)
Quota acquisition × 15-20.9m	100.464***	0.493***	1375.125***	0.488***
	(8.898)	(0.058)	(116.103)	(0.061)
Quota acquisition × 11-14.9m	37.155***	0.408***	362.610***	0.389***
	(2.992)	(0.028)	(41.184)	(0.027)
Panel E: pooled LATE using IV DID FE				
Quota acquisition	44.552***	0.396***	610.531***	0.372***
	(5.803)	(0.120)	(93.497)	(0.116)
Kleibergen-Paap rk Wald F	131.991	131.991	131.991	131.991
Panel F: LATE by license group, using IV DID FE				
Quota acquisition × 21-27.9m	159.263***	0.916***	2695.583***	0.995***
	(19.650)	(0.183)	(338.923)	(0.206)
Quota acquisition × 15-20.9m	122.392***	0.803***	1961.827***	0.809***
	(15.604)	(0.211)	(264.627)	(0.220)
Quota acquisition × 11-14.9m	24.509***	0.298**	253.784***	0.261**
	(6.440)	(0.125)	(77.678)	(0.120)
Kleibergen-Paap rk Wald F	13.406	13.406	13.406	13.406
Observations	30,776	30,776	30,776	30,776

Note: Panels represent specifications. Columns represent dependent variables. All specifications use yearly observations and include year fixed effects, vessel fixed effects, and owner fixed effects. Panel A and B estimate ITT of the trading policy using DID specifications in regression (23). Panel C and D estimate ATT of quota acquisition using DID with fixed effects specifications in regression (24). Panels E and F estimate ATT of quota acquisition using IV DID with fixed effects specification; the main regression is equation (24) but the treatment *Quota acquisition*$_{it}$ is instrumented by the policy assignment *Trade qualified*$_{it}$. Standard errors in parentheses are clustered by vessel's municipality. Significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(a) logged quantity (tonne) (b) logged revenue



(c) logged quantity (tonne) (d) logged revenue



(e) logged quantity (tonne) (f) logged revenue

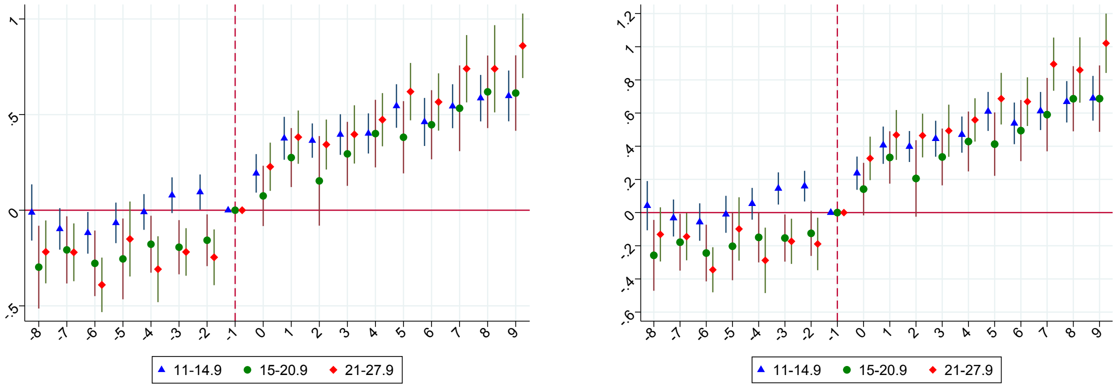


Figure 5: Event studies of ITT of trading policies and ATT of quota acquisition on catch quantity growth and revenue growth

Note: Panels A and B plot the coefficients of the regression (23), except the indicator *Trading Qualified* is interacted with dummies for years before and after the program started. The base group is the licensed length below 11m and year 2004. Panels C–F plot the coefficients of the regression (24), except the indicator *Quota Acquisition*$_{it}$ is interacted with dummies for years before and after the acquisition. The year prior to the quota acquisition year is normalized.

For years after quota acquisition time, we see an immediate effect of quota acquisition. Vessels acquiring quota immediately earns 20% higher than the preceding year.

These increases in catch quantity and revenues are not surprising, because one will expect vessels in the trade qualified group purchase fishing quotas from others to boost the harvest. I now examine the factors behind the harvest boost. Specifically, I first explore the changes in production factors (vessel length, crew size, fishing distance, fishing frequencies). I then investigate the change in productivity.

6.2 Transacted Fish Sales Prices

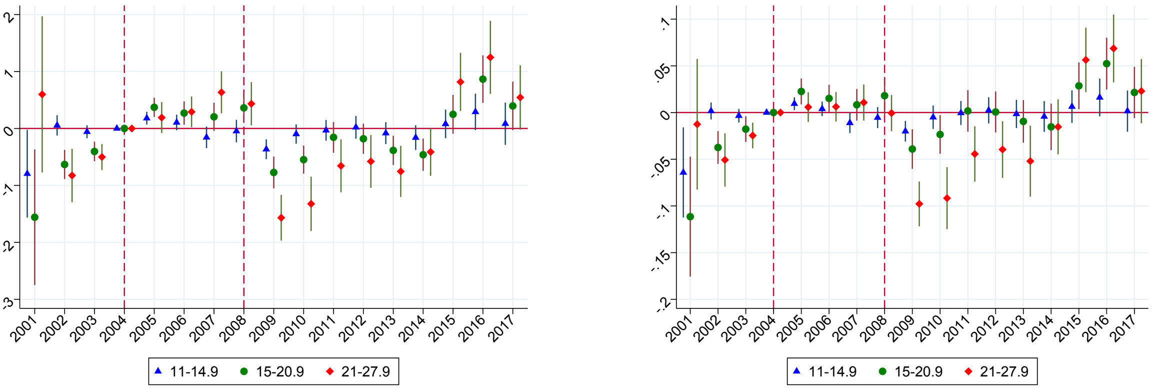
In theory, the concern for the abuse of market power on fish sales prices is attributed to two reasons. First, fewer fishers can reduce the competition of the market, thereby offering market power for those that stay in the market. Second, fishing firms that acquire quotas can increase their sizes and have the capability to negotiate with fish buyers. Investigating the specific effects of each channel is beyond the scope of this paper. I, instead, test the overall effect of the consolidation on transacted prices. Note that although both channels imply positive effects on prices, the consolidation may not induce a price increase. The reason is the trading program imposes the maximum amount of additional quotas a vessel may acquire (quota ceiling). This quota ceiling helps mitigate the market power abuse, as aiming to prevent consolidation as the policy goal. The ceiling would be effective to achieve that goal if there were a substantial number of vessels after the trading, keeping the landing market competitive on the seller side.

In the empirical design, the difference-in-difference strategy may underestimate the effects of consolidation on prices because of the spillover threats. The trading program forced the acquired vessels to exit the fishery, reducing the number of vessels in the whole fishery. At the same time, vessels less than 15 meters are subject to the decommissioning policy during 2004–2008, which reduces the number of vessels that are less than 15 meters. Hence, the non-tradable group that has licensed length 0–10.9m and mostly has actual length less than 11 meters might experience a unit price gain from 2004, causing the DID estimator to underestimate the effect of consolidation on prices.

Results in Table F4 in the appendix show either insignificant or very weak effects of consolidation on prices. ITTs are positive for the two biggest licensed length groups whereas ATTs are around zero. Figure 6 that plots event studies for fish sales prices reveals trip-level transacted fish prices follow stable trends over years before and after the quota acquisition month. Hence, the DID strategy shows no evidence for a change in prices due to the trading program and the quota acquisition. Furthermore, Figure 7 indicates the Herfindahl-Hirschman Index, despite the increasing consolidation over years in the two upper licensed length groups, is still very low (less than 0.03). The fishing market remains very competitive. I conclude that the quota trading policy and quota acquisition did not affect fish sales prices at all.

(a) fish sales price (NOK/kg)

(b) logged fish sales price



(c) fish sales price (NOK/kg)

(d) logged fish sales price

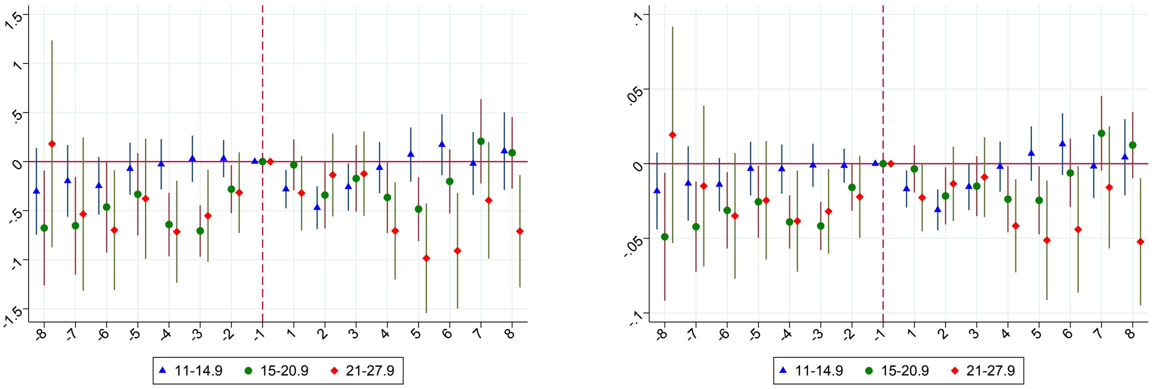


Figure 6: Event studies of ITT of trading policies and ATT of quota acquisition on trip-level transacted fish sales price

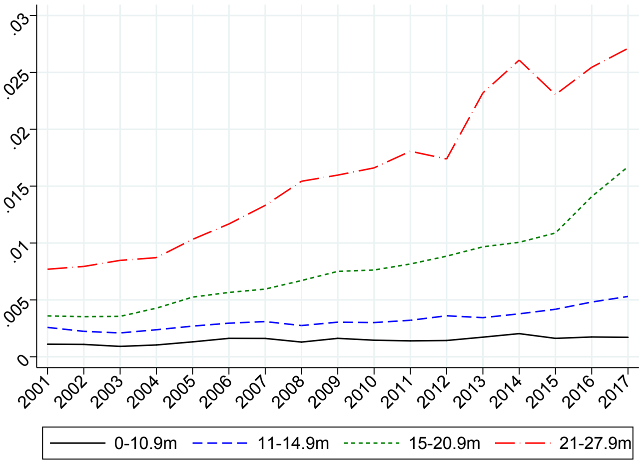
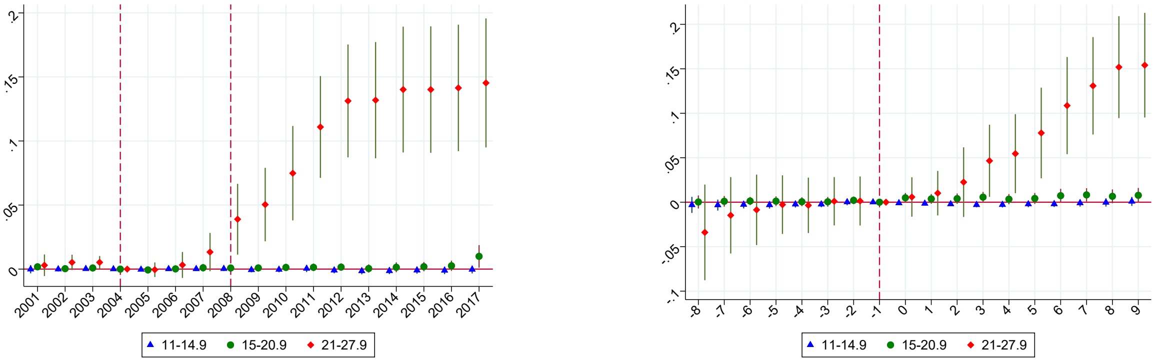


Figure 7: Change in Herfindahl-Hirschman Index over years by licensed length group

6.3 Production Factors

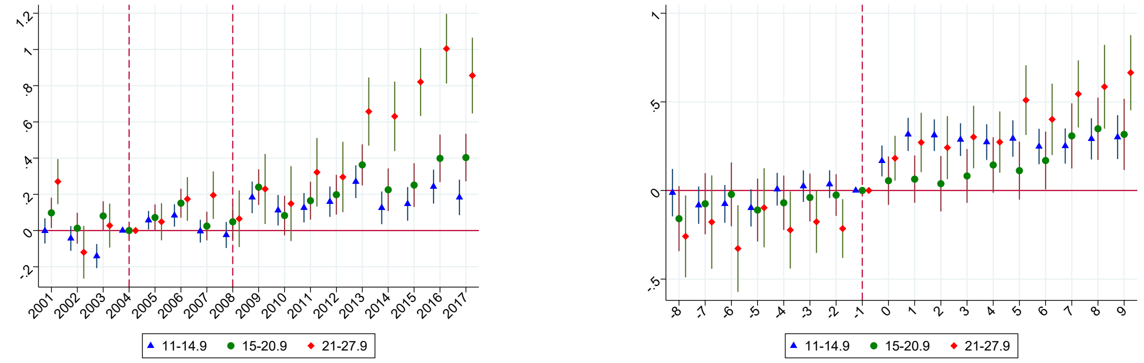
To make use of the higher quota after the trading rights and acquisition, fishing firms must have expanded their production inputs to boost harvest in a year. Table F2 in Appendix F reports the ITT, ATT, and LATE estimates of the effects on vessel length, crew size, fishing distance (distance from the fishers' municipality to major catch location), and the number of trips in a year. Figure 8 plots the corresponding event studies. Among the four production factors, we see the clear and significant increases in vessel length and the number of fishing trips. The trading policy seems to reverse the decreasing trend in crew size, and may imply a slight increase in labor usage. However, compared to the patterns in vessel length and the number of trip, I conclude the policy did not lead to investment in labor. Neither does it for fishing distance.

Hence, significant expansion in production factors comes from the changes in vessel lengths and the trip numbers. Because different licensed length groups are subject to different quota ceilings, the investment in production factors follows a heterogeneous pattern for different license groups. Whereas the biggest licensed length group (21–27.9m) invested in 9% longer vessels and went fishing more frequently (by 24%–46%), the other two tradable licensed length group (11–14.9m and 15–20.9m) did not expand their vessels and instead, only went fishing more frequently to utilize larger fishing quotas. Table F2 even reveals vessels in the licensed length group of 11–14.9m reduced their lengths by 1%. A reason may be the fjord fishing regulation starting from 2004 that granted the fishing rights within fjord areas for only vessels less than 15 meters (actual length). Thus, once the quota ceiling on the tradable group of 11–14.9m was low enough, vessels in this group opted to be strictly less than 15 meters to take an advantage of the fjord fishing rights and go fishing more frequently in the fjord areas to fill up the modest quotas.



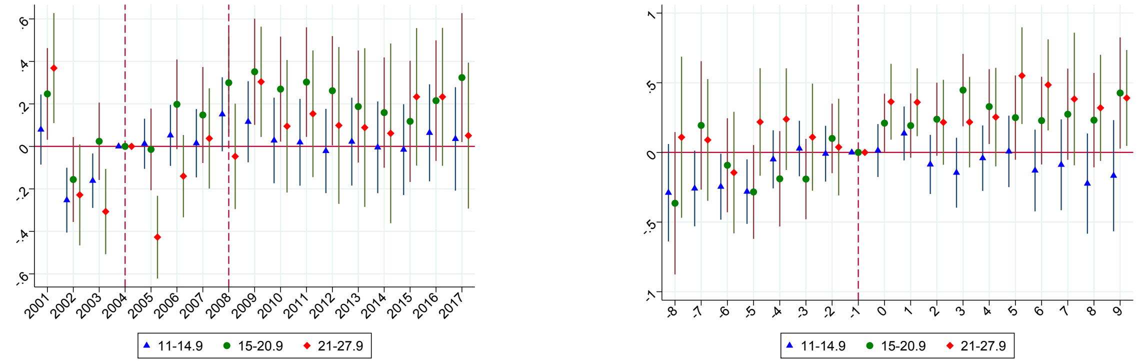
(a) logged length

(b) logged length



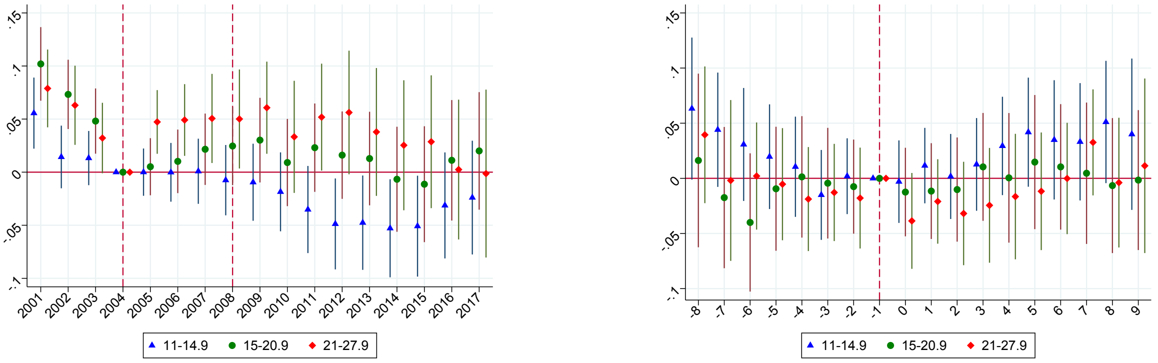
(c) logged # trips

(d) logged # trips



(e) logged distance

(f) logged distance



(g) logged crew

(h) logged crew

Figure 8: Event studies of ITT of trading policies and ATT of quota acquisition on inputs

6.4 Productivity

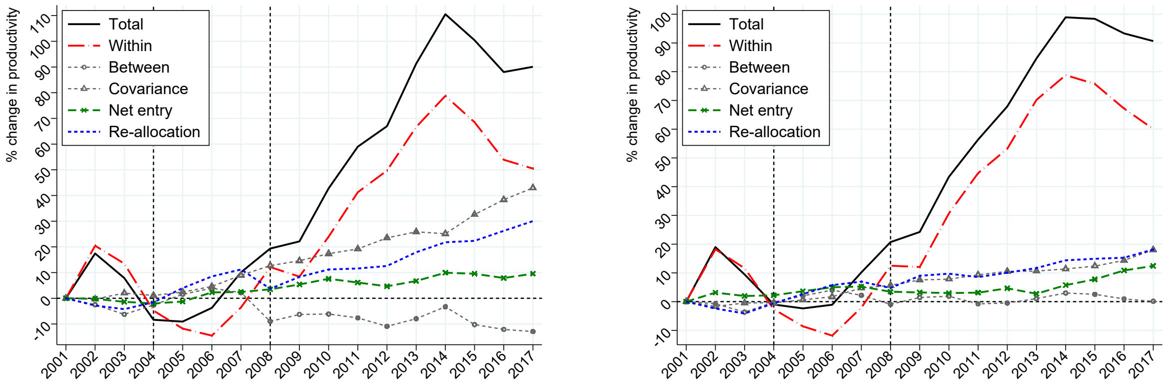
Before exploring the change in vessel-level productivity and its connection to the cap-and-trade policy, I first check the dynamics of the group productivity to examine the roles of reallocation in output shares and net entry. As suggested by Baily et al. (1992); Haltiwanger (1997); De Loecker and Syverson (2021), changes in productivity are attributed to three factors: within-firm factors (such as unobservable input quality, intangible capital, managerial practices, and organization structure), reallocation in market shares, and changes in market structure (net entry). Because the trading-quota policy in the Norwegian cod fishery directly affects the vessel outputs and the number of vessels in the fishery, examining the roles of share reallocation and net entries is important before investigating the vessel-level productivity in the difference-in-difference framework. To do that, I follow Haltiwanger (1997) and De Loecker, Eeckhout and Unger (2020) to decompose the change in group productivity into four components: Δwithin, Δbetween, Δcovariance (Δcross term), and Δnet entry. That is,

$$\begin{aligned}
\Delta\omega_t &\equiv \sum_i s_{it}\omega_{i,t} - \sum_i s_{i,t-1}\omega_{i,t-1} \\
&= \underbrace{\sum_{\text{continuers } i} s_{i,t-1}\Delta\omega_{i,t}}_{\Delta\text{ within}} + \underbrace{\underbrace{\sum_{\text{continuers } i} \Delta s_{i,t}\tilde{\omega}_{i,t-1}}_{\Delta\text{between}} + \underbrace{\sum_{\text{continuers } i} \Delta s_{i,t}\Delta\omega_{i,t}}_{\Delta\text{cross term}}}_{\Delta\text{reallocation}} \\
&\quad + \underbrace{\sum_{\text{entrants } i} s_{i,t}\tilde{\omega}_{i,t} - \sum_{\text{exiters } i} s_{i,t-1}\tilde{\omega}_{i,t-1}}_{\Delta\text{net entry}} .
\end{aligned} \tag{31}$$

In the formula, $\Delta\omega_t$ is the change in group-level productivity between period $t-1$ and period t. The group-level productivity in a period, $\omega_{i,t}$, is the average vessel-level productivity weighted by the vessel share of (licensed length) group output. Note that while $\Delta\omega_{i,t}$ is the change in vessel-level productivity during the period, $\tilde{\omega}_{i,t}$ is the deviation between the vessel productivity from the group index at t, $\tilde{\omega}_{i,t} = \omega_{i,t} - \omega_t$. So, the first term Δwithin measures the average change merely due to a change in within-vessel productivity, given the market shares unchanged. The second term Δbetween measures the change due to an increase in market share while keeping the productivity fixed. If this term is increasing, it captures that vessels that had higher productivity at the first time are increasing their market shares. The Δcross term measures the joint change in market share and productivity. The Δbetween and Δcross term, together, captures the change due to reallocation of market shares across vessels in the market. Finally, the term Δnet entry measures the change due to market structure.

Figure 9 plots the decomposition by year, that is, the decomposition of the change in productivity from 2001 to each year. Productivity generally increases during 2001–2017, starting in 2006. Tables G5 shows the decomposition numbers in selected periods. For all four licensed length groups, the within effect contributes most to the change in productivity, accounting for more than 70% of the productivity improvement since 2008. The net entry effect and re-allocation effect contribute modestly to the change in productivity, but their contributions are increasing, whereas the within effect starts to decrease since 2014. Furthermore, the net entry effect and the re-allocation effect have different contributions when comparing the change in the non-tradable license group (licensed length 0–10.9m) with the change in the tradable groups. In the non-tradable license group, the re-allocation effect dominates the net entry effect. In contrast, the net entry effect appears to outweigh the re-allocation effect for tradable license groups. The difference in the effect pattern between those groups may be resulted from the fact that the trading policy requires vessels exit after selling quotas.

(a) licensed length 0–10.9m (b) licensed length 11–14.9m



(c) licensed length 15–20.9m (d) licensed length 21–27.9m

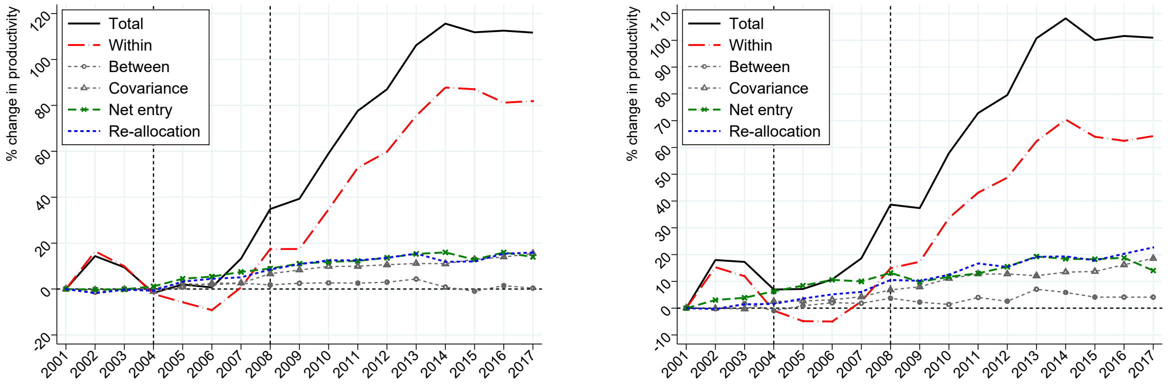


Figure 9: Decomposition of change in productivity by year

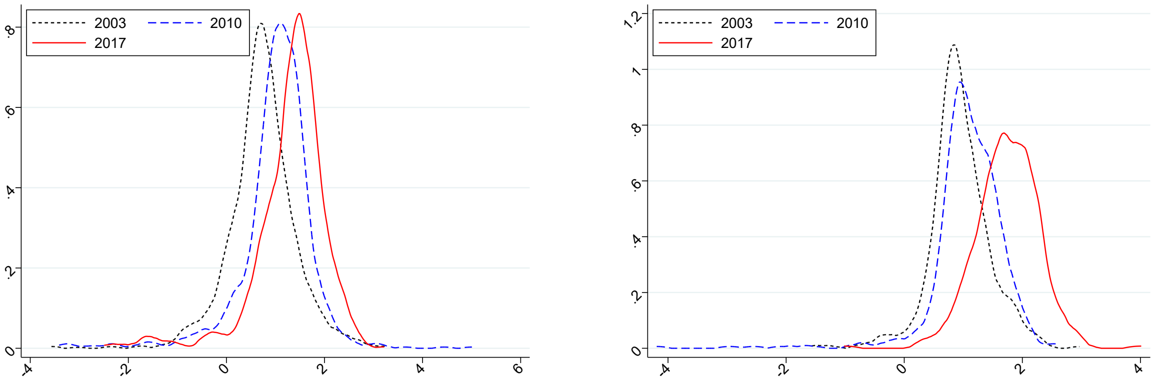
Note: All changes relative to 2001. The production function is estimated using the proxy variable approach.

I now look into the change in productivity at vessel level. Figure 10 plots the distribution of productivity by licensed length group and year. The figure plots the productivity index estimated using the proxy variable approach.[15] Over years, productivity in all of the four groups has shifted to the right side. The shifts to the right are firstly attributed to the common improvement in productivity over time. One instant reason is technology in harvest may advance over

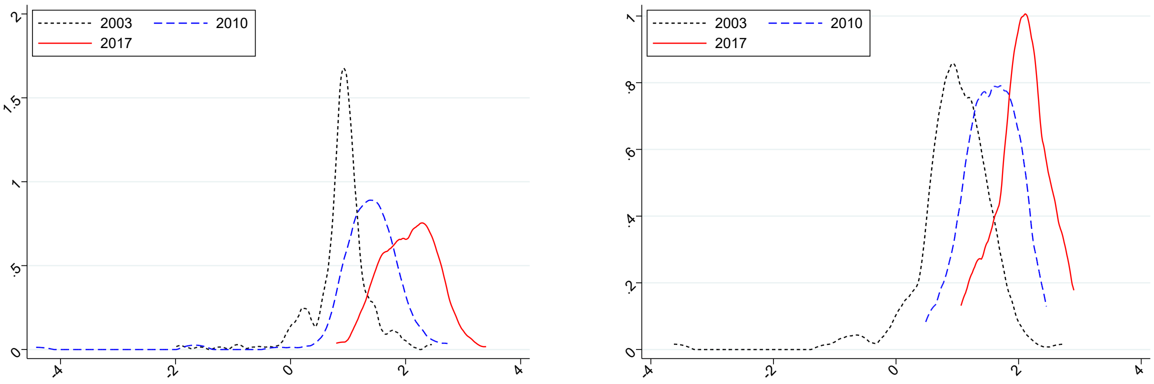
---

[15] Appendix E shows the distribution of productivity index estimated by the OLS with fixed effects and the dynamic panel approaches.

time. Another reason is fisherman may adapt better to weather conditions and obtain productivity gains over time. Furthermore, although only the three upper license group has the trading program, spillovers of the trading impact on the whole fishery biomass may increase common fish stocks, allowing the regulator to increase total allowable catch in all four groups.


(a) productivity in licensed length 0–10.9m

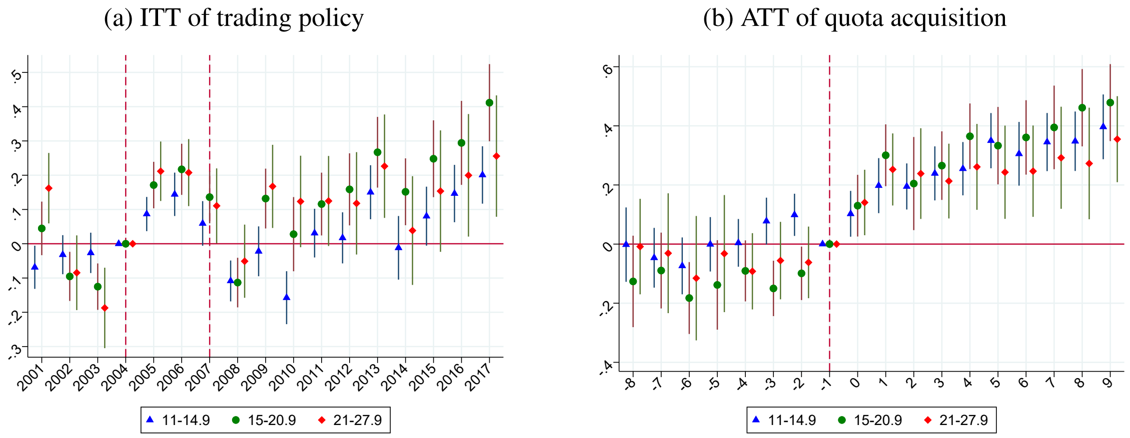
(b) productivity in licensed length 11–14.9m


(c) productivity in licensed length 15–20.9m

(d) productivity in licensed length 21–27.9m

Figure 10: Distribution of productivity (proxy variable approach)

Despite the same direction in the shifts of productivity distribution in all four groups, the evolution of the distribution shows remarkably different patterns between the non-tradable group and the tradable groups. The distribution of productivity in the non-tradable group (0–10.9m) just simply shifts itself over years. Its peak level, variance, and tail length nearly remain the same as before. This implies the change in productivity of vessels in this group is purely induced by the systematic shocks over time. On the other hand, the distributions of productivity in the other tradable groups shift and change their shapes. The variance has decreased and more importantly, the lower tail has considerably shortened over time. In 2010, the distributions of productivity in the two middle groups (11–14.9m and 15–20.9m) had a long left tail, suggesting these two groups had a few vessels with extremely low productivity relative to the other vessels in the group. In 2017, these left tails were cut, and the new distributions even did not appear skewed. The shortening in the tails implies that the lowest productive vessels has exited from the fishery. For vessels in the license group 21–27.9m, exits of low productive vessels has


(a) ITT of trading policy

(b) ATT of quota acquisition

Figure 11: Impacts of the trading scheme and quota acquisition on productivity

Note: Productivity is estimated using proxy variable approach. Panel A plots the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panel B plots the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.

happened sooner than the two middle groups. Table G6 in the appendix shows average vessel productivity index by trading and entry/exit status. The table confirms that exiters are the lowest productive, lower than the stayers that never buy quotas. Entrants have high productivity when entering the industry, even higher than the stayers that acquire additional quotas, in some periods.

Finally, I run difference-in-difference regressions on vessel productivity to examine the effects of trading policy on within-vessel productivity. Table F3 in Appendix F reports the ITT, ATT, and LATE estimates of the policy impacts on productivity. Figure 11 plots the event-study coefficients. In general, we do not see the program has a significant effect on all individuals in the trading group. Instead, the trading program has substantial impact only on vessels that do acquire quotas. Acquiring additional quotas help vessel increase its productivity by 20%–40% on average.

6.5 Economies of Scale and Production Cost

I now investigate how the cap-and-trade affects production cost. Although I do not observe the cost, I have shown that variation in output and input can infer output elasticity of cost and how this elasticity infers economies of scale and change in average cost per unit.

Table 4 contrasts the means and ranges of output elasticities of total costs between before and after the trading policy. The licensed length group 0–10.9m is not allowed to trade, but we contrasts its range of cost elasticity in the 2001–2007 period to the level in the 2009–2017 period. The table reports the cost elasticity using the proxy variable estimator for the production function. Appendix E reports results using other approaches for the production function estimation. The first important point of the results is that the output elasticities of total costs in the pre-trade program are less than 1. This suggests vessels were having economies of scale before the trading program. The second point to notice is within each license group the mean and the max levels become higher in the post-trade period. These increases reveal that vessels are moving along the average cost curves starting on the left side of the curves, thereby suggesting the average costs were going down toward minimum levels of average cost.

Table 4: Summary statistics of estimates of cost indices by licensed length group

	Pre-trade-program					Post-trade-program				
	count	mean	sd	min	max	count	mean	sd	min	max
Panel A: Output elasticity of total costs, using the proxy variable estimator for the production function										
0–10.9m	8,470	0.372	0.049	0.203	0.602	8,362	0.377	0.060	0.203	0.849
11–14.9m	3,590	0.433	0.061	0.244	0.677	3,637	0.449	0.074	0.220	0.827
15–20.9m	995	0.458	0.066	0.249	0.703	2,367	0.514	0.089	0.278	0.918
21–27.9m	433	0.456	0.068	0.289	0.738	1,016	0.524	0.097	0.313	1.014

Note: Pre-trade-program period and post-trade period for licensed groups 15–20.9m and 21–27.9m are 2001–2003 and 2005–2017. For licensed length group 11–14.9m, they are 2001–2007 and 2009–2017. Licensed length group 0–10.9m is not allowed to trade during 2001 and 2017, but we compare period 2001–2007 to period 2009–2017.

Figure 12 examines the change in cost elasticity over time within a license group by exploring the distribution of the cost elasticity over years by licensed length group.[16] Two features are noted. First, the distribution of cost elasticity in the lowest group (0–10.9m) almost remains unchanged from 2003 to 2017, suggesting the production costs of vessels in this group nearly do not change. On the other hand, the distributions of cost elasticities in the other three upper groups significantly become fattened and more positively skewed. So, vessels in these groups on average have substantially higher cost elasticities over time. There is also sizable variance in cost elasticity among vessels within each upper group. The second noticeable feature is that it is the upper tail that drives the increase in cost elasticity of vessels in the three upper group. For the two highest upper groups, we also see the left tails have been cut, implying a number of high-average-cost vessels (of which cost elasticities are far much smaller than 1) exits the fishery.

In summary, we see that vessels in the fishery almost have cost elasticity below 1 before 2004. This means they are operating on the side of economies of scale on the average cost curve. The fact that their cost elasticity has increased over time implies that they have expanded their output at lower average cost levels, moving toward the minimum-average-cost operation. To formally test whether the movement is caused by the cap-and-trade program, I examine the change in cost elasticity within a vessel over time using the DID approach. Table 5 and

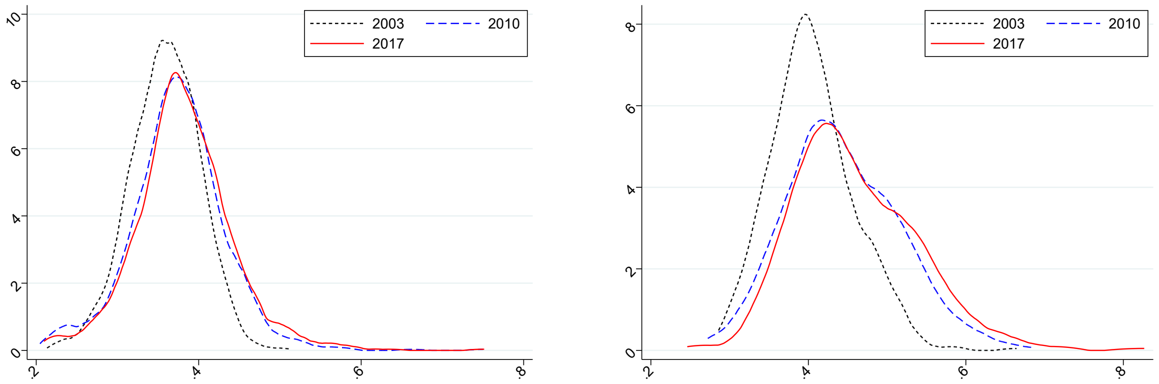
---

[16] These estimated output elasticity of total costs are calculated from the proxy variable approach of estimating a production function. Appendix E shows the distribution that uses the OLS with fixed effects and dynamic panel approaches of the production function.

Figure 13 report the estimates and the event study. I find that the trading program plays an important role in raising the output elasticity of cost, thereby pushing the fishing operation toward the minimum-average-cost operation. The push happened strongest in the highest group (21–27.9m), because this group offers the largest room to acquire additional quota amount, offering the vessels possible biggest move in harvest expansion. Vessels that acquire additional quotas in this group incurred additional 6–8 percentage point in the output elasticity of cost. A back-of-the-envelope calculation uses these estimates finds that a vessel in the licensed group of 11–14.9m, 15–20.9m, and 21–27.9m saves 0.42%, 4.07%, and 15.60% of average cost, respectively.

(a) elasticity of total costs, license 0–10.9m

(b) elasticity of total costs, license 11–14.9m



(c) elasticity of total costs, license 15–20.9m

(d) elasticity of total costs, license 21–27.9m

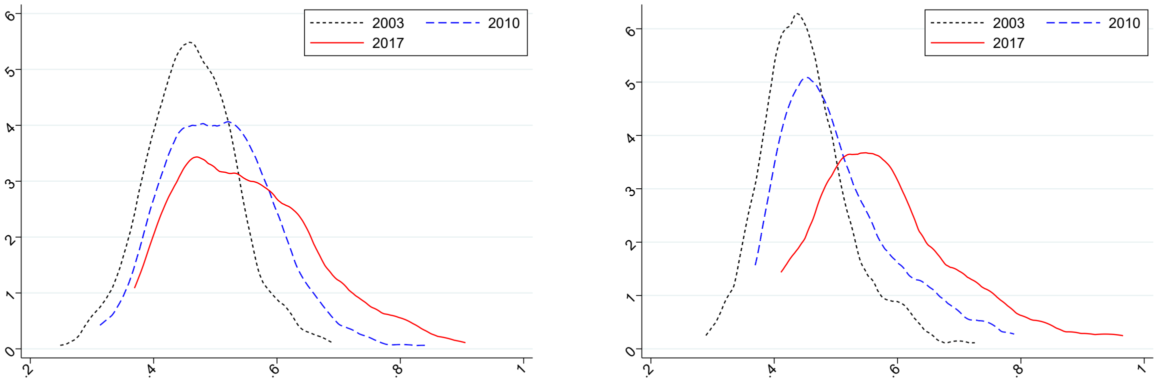


Figure 12: Distribution of output elasticity of total costs (proxy variable approach)

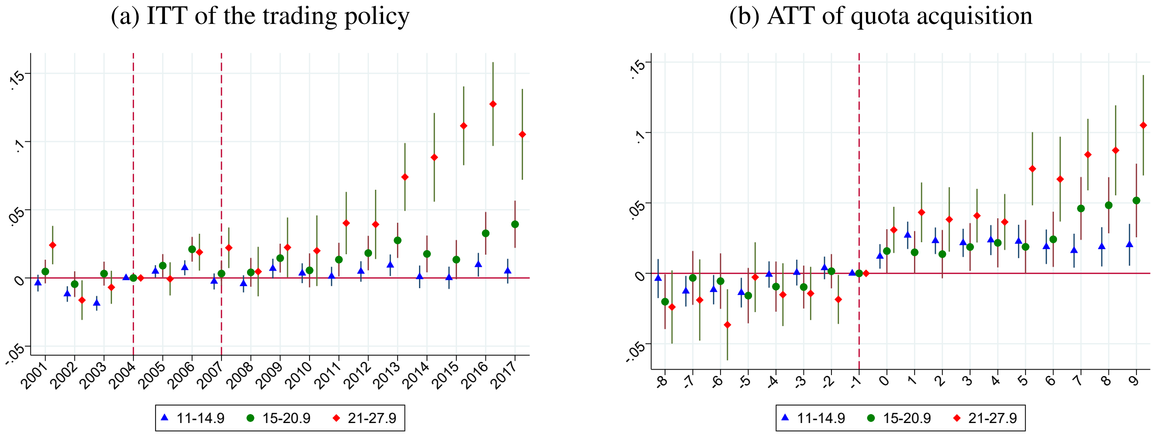
7 Economies of Scale vs. Cost Shifting

I perform the decomposition method in proposition (6) to entire vessels in the fishery to investigate the dynamic of output change ($\ln Q_{it}$) of a vessel by licensed length group. Figure 14 plots the decomposition of the change by year (relative to 2001). The year index is the vessel-change index weighted by a vessel's share of catch in the licensed length group. Out-

Table 5: Effects of trading policy on output elasticity of total costs

	(1)	(2)	(3)
	OLS-with-FE estimator	proxy-variable estimator	dynamic panel estimator
	cost elasticity ϕ	cost elasticity ϕ	cost elasticity ϕ
Panel A: ITT (pooling all trade qualified groups)			
Trade qualified	0.005^{**}	0.004^{**}	0.004^{**}
	(0.002)	(0.002)	(0.002)
Panel B: ITT by trade qualified group (license group)			
21-27.9m × From 2004	0.015^{**}	0.012^{*}	0.013^{**}
	(0.007)	(0.006)	(0.006)
15-20.9m × From 2004	0.006	0.004	0.004
	(0.005)	(0.004)	(0.004)
11-14.9m × From 2008	0.003	0.003	0.002
	(0.002)	(0.002)	(0.002)
Panel C: pooled ATT using DID FE			
Quota acquisition	0.039^{***}	0.033^{***}	0.032^{***}
	(0.003)	(0.003)	(0.003)
Panel D: ATT by license group, using DID FE			
Quota acquisition × 21-27.9m	0.075^{***}	0.062^{***}	0.061^{***}
	(0.008)	(0.007)	(0.007)
Quota acquisition × 15-20.9m	0.040^{***}	0.034^{***}	0.032^{***}
	(0.007)	(0.006)	(0.005)
Quota acquisition × 11-14.9m	0.029^{***}	0.025^{***}	0.023^{***}
	(0.004)	(0.003)	(0.003)
Panel E: pooled LATE using IV DID FE			
Quota acquisition	0.023^{**}	0.019^{**}	0.018^{**}
	(0.010)	(0.008)	(0.008)
Kleibergen-Paap rk Wald F	131.991	131.991	131.991
Panel F: LATE by license group, using IV DID FE			
Quota acquisition × 21-27.9m	0.062^{**}	0.048^{**}	0.051^{**}
	(0.027)	(0.023)	(0.023)
Quota acquisition × 15-20.9m	0.038	0.028	0.026
	(0.031)	(0.027)	(0.025)
Quota acquisition × 11-14.9m	0.018^{**}	0.015^{*}	0.014^{*}
	(0.009)	(0.008)	(0.007)
Kleibergen-Paap rk Wald F	13.406	13.406	13.406
Observations	30,776	30,776	30,776

Note: Panels represent specifications. Columns represent dependent variables. All specifications use yearly observations and include year fixed effects, vessel fixed effects, and owner fixed effects. Panels A and B estimate ITT of the trading policy using DID specifications in regression (23). Panels C and D estimate ATT of quota acquisition using DID with fixed effects specifications in regression (24). Panels E and F estimate ATT of quota acquisition using IV DID with fixed effects specification; the main regression is equation (24) but the treatment *Quota acquisition*$_{it}$ is instrumented by the policy assignment *Trade qualified*$_{it}$. Standard errors in parentheses are clustered by vessel's municipality. Significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.


(a) ITT of the trading policy

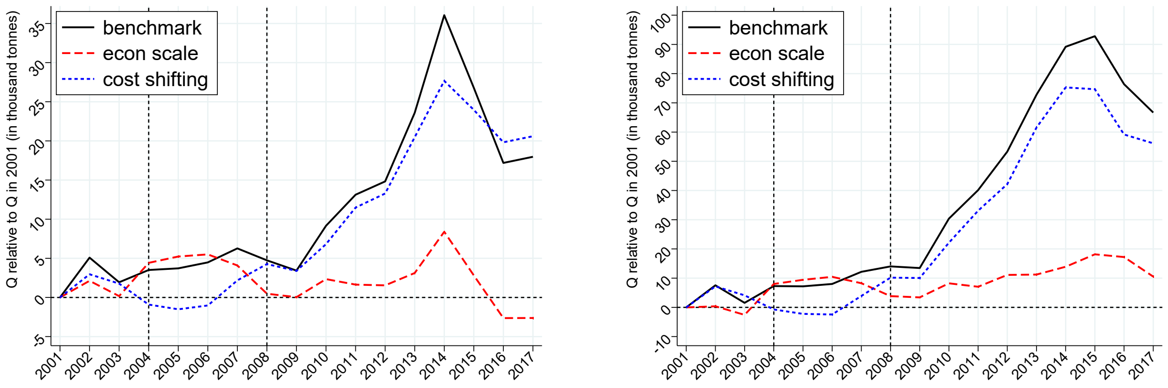
(b) ATT of quota acquisition

Figure 13: Impacts of the trading policy and quota acquisition on output elasticity of total cost
Note: Output elasticity of total cost ϕ is estimated using a proxy variable approach. Panel A plots the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panel B plots the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.

put generally increases over years for all licensed length groups but is underlied by different patterns of the decomposing effects. Vessels in the non-tradable group (0–10.9m) have little economies-of-scale effect. The cost shifting plays the main role in explaining the change in output in this group. For vessels in the tradable group, both the two effects contribute to the change in output, but the relative contribution of the two decomposing effects differs in different periods. Between 2001 and 2003, the cost-shifting effect contributes more to the change in output. From 2004 through 2007, almost only the economies-of-scale effect contributes to the output growth; the cost shifting has no effect at all. Since 2007, the cost-shifting effect dramatically increases and outweighs the economies-of-scale effect, except the vessels in the group 21–27.9m. The output growth in vessels in the biggest licensed length group (21–27.9m) is generally attributed to the econonomies of scale effect. In recent years from 2014, the two effects and the output decrease, except for vessels in the licensed length 15–20.9m.

Figure 15 plots the decomposition of output change relative to the first trading time for vessels that ever acquire tradable quotas. Note that this is the change in output within a vessel relative to its own first trading time and does not take the change in output of other vessels into account. Not surprisingly, vessels increase their catches dramatically after acquiring quotas. However, whereas the economies-of-scale effect does not contribute to the output growth of vessels in the group 11–14.9m, it explains the output growth for vessels in the two upper groups (15–20.9m and 21–27.9m) in the first two years after acquiring quotas. Three years after acquiring quotas, cost shifting plays the main role to increase output of vessels in the group 15–20.9m, but it shares quite an equal contribution to the output growth of vessels in the group 21–27.9m, compared to the role of economies of scale.

Figure 16 shows average vessel-level output and counterfactual outputs due to economies-of-scale effect and cost-shifting effect by licensed length group and trading status. The first experiment (long dash red line) shows the evolution of the output growth as if there were only economies-of-scale component and the other cost-shifting effect were 0. The second experiment (short dash blue line) shows the path of the output growth if the change had been only due to cost shifting. Four patterns emerge. First, whereas output for trading vessels remains quite stable during 2001–2004, it significantly increases since 2004 increase over time. Output for never-trading vessels also increases but after 2009. This suggests there may be spillovers of the quota acquisition policy to never-trading stayers due to improvement in the common stock biomass or the change in market structure (exits of selling-quota vessels). The spillovers can understate the effects of the trading policy on individual vessel performance. However, the trend in the actual benchmark output and the counterfactual single-effect outputs in the pre-trade periods (before 2004 for the groups 15–20.9m and 21–27.9m and before 2008 for the group 11–14.9m) confirms the significant causal impacts of the trading policy on output. Whereas the second pattern in this figure notes the systematic lows in all output benchmark, economies-of-scale output, and cost-shifting output between ever-trading vessels and never-trading vessels, the third pattern highlight the systematic gaps are stable for the pre-policy periods, implying the parallel trend condition is satisfied. Indeed, the fact that both counterfactual economies-of-scale output and cost-shifting output of never-trading vessels are lower than those of ever-trading vessels in the pre-policy periods suggests vessels that choose to acquire quotas are the ones that had low operating costs due to economies of scale and high productivity. However, this selection bias can be controlled by switching to the parallel trend identifying condition. The final interesting pattern is that economies of scale plays an important role in explaining the output growth of vessels in the biggest licensed length group (21–27.9m), although it is dominated by the cost-shifting effect for vessels in the other two tradable groups. Moreover, whereas the economies-of-scale effect does not contribute to the output growth of never-trading vessels in the two middle groups, it is positive and increasing for vessels that do *not* trade in the group 21–27.9m, suggesting these never-trading vessels also reallocate their inputs to exploit economies of scale.



(a) licensed length 0–10.9m (b) licensed length 11–14.9m

(c) licensed length 15–20.9m (d) licensed length 21–27.9m

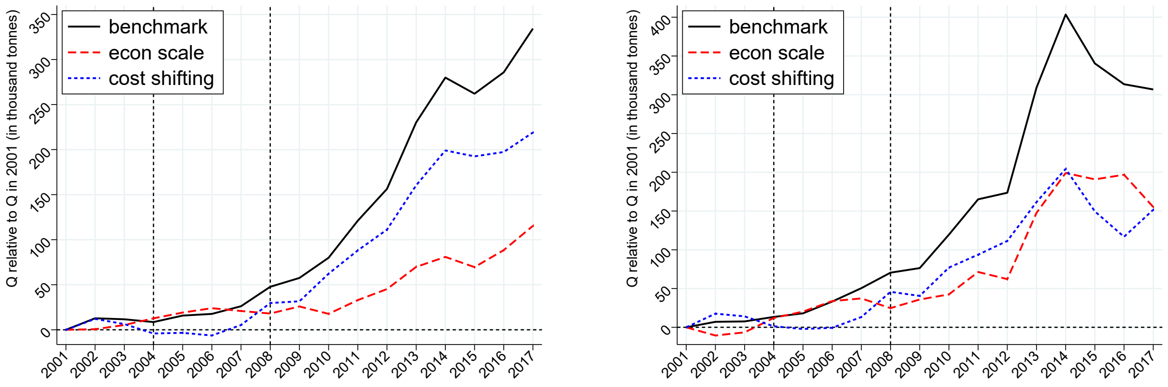
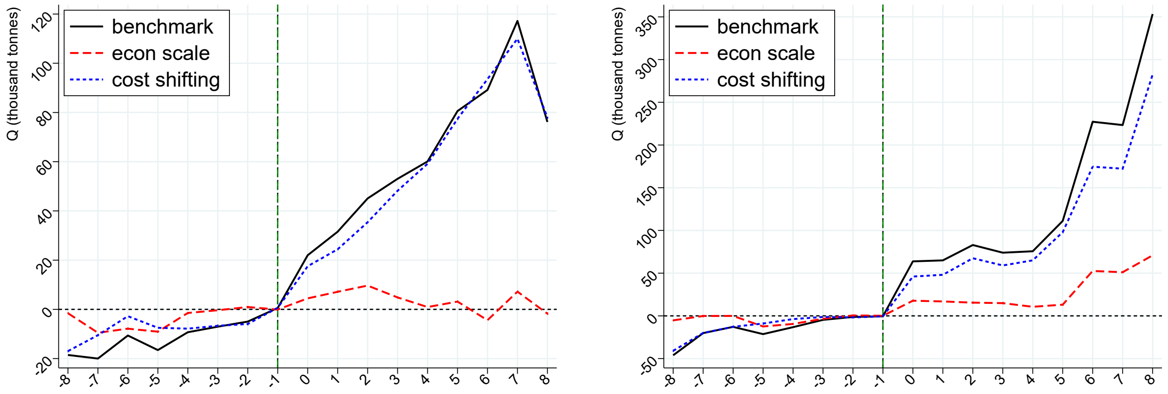


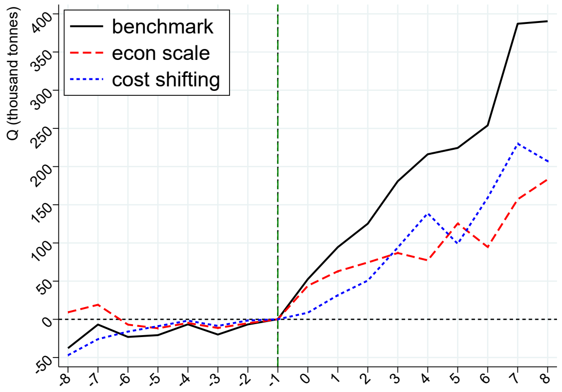
Figure 14: Decomposition of change in Q_{it} (thousand tonnes) by year

Note: The figure plots the change in output within a vessel from 2001 to each year. The average change number for each year is weighted by the vessel share of group catch.



(a) licensed length 11–14.9m

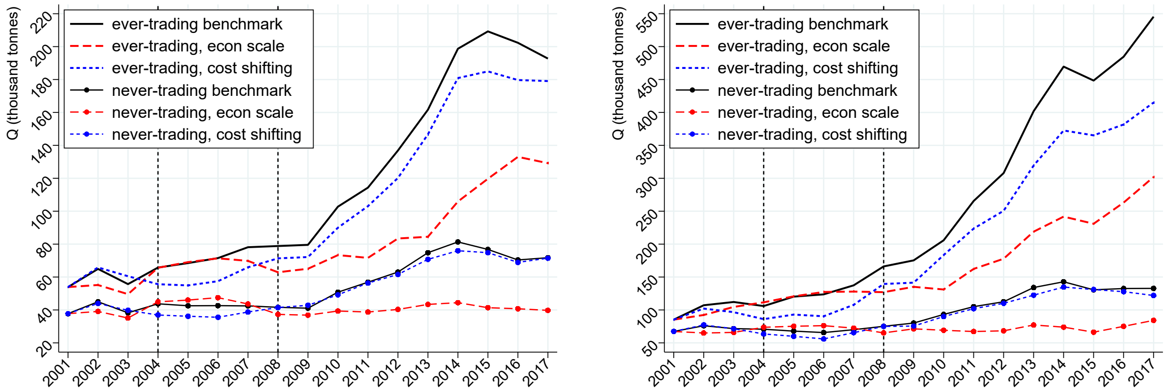
(b) licensed length 15–20.9m



(c) licensed length 21–27.9m

Figure 15: Decomposition of change in Q_{it} (thousand tonnes) by years from the first time a vessel acquires traded quotas

Note: The figure plots the change in output within a vessel over years. All changes are relative to the year when a vessel acquires traded quotas in its first time.



(a) licensed length 11–14.9m

(b) licensed length 15–20.9m

(c) licensed length 21–27.9m

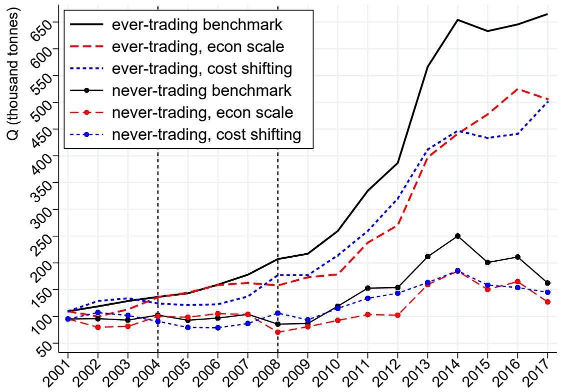


Figure 16: Counterfactual catch by trading action and licensed length group

Note: The figure plots the counterfactual output (thousand tonnes) by quota-acquisition status. The curves with (without) dot markers represents output growths of vessels that never (ever) acquired quotas. The benchmark represents the average output (Q_{it}) weighted by a vessel's share of catch in the licensed length group. The long dash red line shows the evolution of the output as if there were only "economies of scale" effect and the other "cost shifting" effect were 0. The short dash blue line shows the path of the output if the change had been only due to cost shifting.

8 Conclusion

This paper studies the impact of cap-and-trade program, relative to command-and-control instrument, on individual performance in the main product market. I specifically look at the policy performance in fishery context because of two reasons. First, there are little empirical studies that study market-based instruments vis à vis prescriptive regulation in resource economics. Second, the cap regulation in resource economics is directly imposed on the main product the firm produces rather than by-products such as emissions. This direct constraint will have apparent impact on the firm performance in the main product market.

After demonstrating the consolidation effects of cap-and-trade on the standard performance outcomes such as output and revenue, I investigate the policy impacts on the pricing side and the production cost side. Because production cost is not observed, I introduce a method to estimate the output elasticity of total costs and economies of scale using data on output and input quantity. This method relies on the cost-minimization behavior assumption and is applicable beyond fishery context, offering a tool for an empirical economist to review the estimate of a production function and infer production costs with data on output and input.

Results using a difference-in-difference approach to compare the change in output elasticity of cost from pre-trade period to post-trade period between vessels in tradable group and in non-tradable group show the cap-and-trade policy has caused the output elasticity of cost to increase. This evidence of the increase together with the finding that cost elasticity was less than 1 before the policy imply vessels had economies of scale before the policy and trading has induced vessels to expand output at lower average cost of a unit, suggesting the cap-and-trade has improved the cost efficiency in the fishery operation compared to the prescriptive nontradable cap instrument.

I further propose a decomposition method to distinguish the output growth due to cost efficiency (economies of scale) from the growth due to cost effectiveness (cost shifting or changes in productivity). I find productivity improvement induced by trading did not contribute much to output growth in the first three years after acquiring quota. It is economies of scale that plays the main role in explaining the output growth at initial stages.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. “Identification properties of recent production function estimators.” *Econometrica*, 83(6): 2411–2451.

Ang, Beng Wah, and Fang Q Zhang. 2000. “A survey of index decomposition analysis in energy and environmental studies.” *Energy*, 25(12): 1149–1176.

Arellano, Manuel, and Olympia Bover. 1995. “Another look at the instrumental variable estimation of error-components models.” *Journal of Econometrics*, 68(1): 29–51.

Arellano, Manuel, and Stephen Bond. 1991. “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.” *Review of Economic Studies*, 58(2): 277–297.

Armstrong, Claire W, and Derek J Clark. 1997. “Just fishing? Equity and efficiency in fisheries management regimes.” *Marine Resource Economics*, 12(3): 203–220.

Armstrong, Claire W, Arne Eide, Ola Flaaten, Knut Heen, and Inga Wigdahl Kaspersen. 2014. “Rebuilding the Northeast Arctic cod fisheries–economic and social issues.” *Arctic Review*, 5(1).

Asche, Frank, Håkan Eggert, Eyjolfur Gudmundsson, Ayoe Hoff, and Sean Pascoe. 2008. “Fisher’s behaviour with individual vessel quotas—Over-capacity and potential rent: Five case studies.” *Marine Policy*, 32(6): 920–927.

Baily, Martin Neil, Charles Hulten, David Campbell, Timothy Bresnahan, and Richard E Caves. 1992. “Productivity dynamics in manufacturing plants.” *Brookings papers on economic activity. Microeconomics*, 1992: 187–267.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. “Automobile prices in market equilibrium.” *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, Steven T. 1994. “Estimating discrete-choice models of product differentiation.” *The RAND Journal of Economics*, 242–262.

Blundell, Richard, and Stephen Bond. 1998. “Initial conditions and moment restrictions in dynamic panel data models.” *Journal of Econometrics*, 87(1): 115–143.

Blundell, Richard, and Stephen Bond. 2000. “GMM estimation with persistent panel data: an application to production functions.” *Econometric Reviews*, 19(3): 321–340.

Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson. 2015. “Acquisitions, productivity, and profitability: Evidence from the Japanese cotton spinning industry.” *American Economic Review*, 105(7): 2086–2119.

Bresnahan, Timothy F. 1989. “Empirical studies of industries with market power.” *Handbook of industrial organization*, 2: 1011–1057.

Callaway, Brantly, and Pedro HC Sant’Anna. 2020. “Difference-in-differences with multiple time periods.” *Journal of Econometrics*.

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L Palmer. 2000. “Sulfur dioxide control by electric utilities: What are the gains from trade?” *Journal of Political Economy*, 108(6): 1292–1326.

Chan, H Ron, B Andrew Chupp, Maureen L Cropper, and Nicholas Z Muller. 2018. "The impact of trading on the costs and benefits of the Acid Rain Program." *Journal of Environmental Economics and Management*, 88: 180–209.

Chu, Cindy. 2009. "Thirty years later: the global growth of ITQs and their influence on stock status in marine fisheries." *Fish and Fisheries*, 10(2): 217–230.

Costello, Christopher, John Lynham, Sarah E Lester, and Steven D Gaines. 2010. "Economic Incentives and Global Fisheries Sustainability." *Annual Review of Resource Economics*, 2(1): 299–318.

Costello, Christopher, Steven D Gaines, and John Lynham. 2008. "Can catch shares prevent fisheries collapse?" *Science*, 321(5896): 1678–1681.

De Loecker, Jan. 2013. "Detecting learning by exporting." *American Economic Journal: Microeconomics*, 5(3): 1–21.

De Loecker, Jan, and Chad Syverson. 2021. "An industrial organization perspective on productivity." In *Handbook of Industrial Organization*. Vol. 4, 141–223. Elsevier.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and firm-level export status." *American Economic Review*, 102(6): 2437–71.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The rise of market power and the macroeconomic implications." *Quarterly Journal of Economics*, 135(2): 561–644.

De Loecker, Jan, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik. 2016. "Prices, markups, and trade reform." *Econometrica*, 84(2): 445–510.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition methods in economics." In *Handbook of labor economics*. Vol. 4, 1–102. Elsevier.

Fowlie, Meredith, Stephen P Holland, and Erin T Mansur. 2012. "What do emissions markets deliver and to whom? Evidence from Southern California's NOx trading program." *American Economic Review*, 102(2): 965–93.

Fox, Kevin J, R Quentin Grafton, James Kirkley, and Dale Squires. 2003. "Property rights in a fishery: regulatory change and firm performance." *Journal of Environmental Economics and Management*, 46(1): 156–177.

Frank, Kenneth T, Brian Petrie, Jae S Choi, and William C Leggett. 2005. "Trophic cascades in a formerly cod-dominated ecosystem." *Science*, 308(5728): 1621–1623.

Gandhi, Amit, Salvador Navarro, and David A Rivers. 2020. "On the identification of gross output production functions." *Journal of Political Economy*, 128(8): 2973–3016.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*.

Grafton, R Quentin, Dale Squires, and Kevin J Fox. 2000. "Private property and economic efficiency: a study of a common-pool resource." *The Journal of Law and Economics*, 43(2): 679–714.

Grainger, Corbett A, and Dominic P Parker. 2013. "The Political Economy of Fishery Reform." *Annual Review of Resource Economics*, 5(1): 369–386.

Haltiwanger, John. 1997. "Measuring and analyzing aggregate fluctuations: the importance of building from microeconomic evidence." *Federal Reserve Bank of St. Louis Review*, , (May): 55–78.

Holland, Stephen P, Erin T Mansur, Nicholas Z Muller, and Andrew J Yates. 2020. "Decompositions and policy consequences of an extraordinary decline in air pollution from electricity generation." *American Economic Journal: Economic Policy*, 12(4): 244–74.

Isaksen, Elisabeth Thuestad, and Andries Richter. 2019. "Tragedy, property rights, and the commons: Investigating the causal relationship from institutions to ecosystem collapse." *Journal of the Association of Environmental and Resource Economists*, 6(4): 741–781.

Levinsohn, James, and Amil Petrin. 2003. "Estimating production functions using inputs to control for unobservables." *Review of Economic Studies*, 70(2): 317–341.

Nærings- og fiskeridepartementet. 2003. "Strukturtiltak i kystfiskeflåten (Structural measures in the coastal fishing fleet)." Nærings- og fiskeridepartementet (Ministry of Trade and Industry). Report No. 20 to Parliament (in Norwegian).

Nærings- og fiskeridepartementet. 2006. "Strukturvirkemidler i fiskeflåten (Structural instruments in the fishing fleet)." Nærings- og fiskeridepartementet (Ministry of Trade and Industry). Public Report No. 16 (in Norwegian).

Nærings- og fiskeridepartementet. 2007. "Strukturpolitikk for fiskeflåten (Structural policy for the fishing fleet)." Nærings- og fiskeridepartementet (Ministry of Trade and Industry). Report No. 21 to Parliament (in Norwegian).

Nærings- og fiskeridepartementet. 2016. "Et fremtidsrettet kvotesystem (A future-oriented quota system)." Nærings- og fiskeridepartementet (Ministry of Trade and Industry). Public Report No. 26 (in Norwegian).

Nærings- og fiskeridepartementet. 2019. "Et kvotesystem for økt verdiskaping – En fremtidsrettet fiskerinæring (A quota system for increased value creation – A future-oriented fishing industry)." Nærings- og fiskeridepartementet (Ministry of Trade and Industry). Report No. 32 to Parliament (in Norwegian).

Newell, Richard G, James N Sanchirico, and Suzi Kerr. 2005. “Fishing quota markets.” *Journal of environmental economics and management*, 49(3): 437–462.

OECD. 2011. “OECD Fisheries Policy Reform: National Experiences.”

OECD. 2013. “OECD Review of Fisheries: Policies and Summary Statistics.”

Olley, G Steven, and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica*, 64(6): 1263–1297.

Petrick, Sebastian, and Ulrich J Wagner. 2014. “The impact of carbon trading on industry: Evidence from German manufacturing firms.” *Available at SSRN 2389800*.

Rubens, Michael. 2021. “Market Structure, Oligopsony Power, and Productivity.” SSRN.

Standal, Dag, and Bernt Aarset. 2008. “The IVQ regime in Norway: A stable alternative to an ITQ regime?” *Marine Policy*, 32(4): 663–668.

Standal, Dag, and Frank Asche. 2018. “Hesitant reforms: The Norwegian approach towards ITQ’s.” *Marine Policy*, 88: 58–63.

Stavins, Robert N. 2011. “The problem of the commons: still unsettled after 100 years.” *American Economic Review*, 101(1): 81–108.

Stiebale, Joel, and Dev Vencappa. 2018. “Acquisitions, markups, efficiency, and product quality: Evidence from India.” *Journal of International Economics*, 112: 70–87.

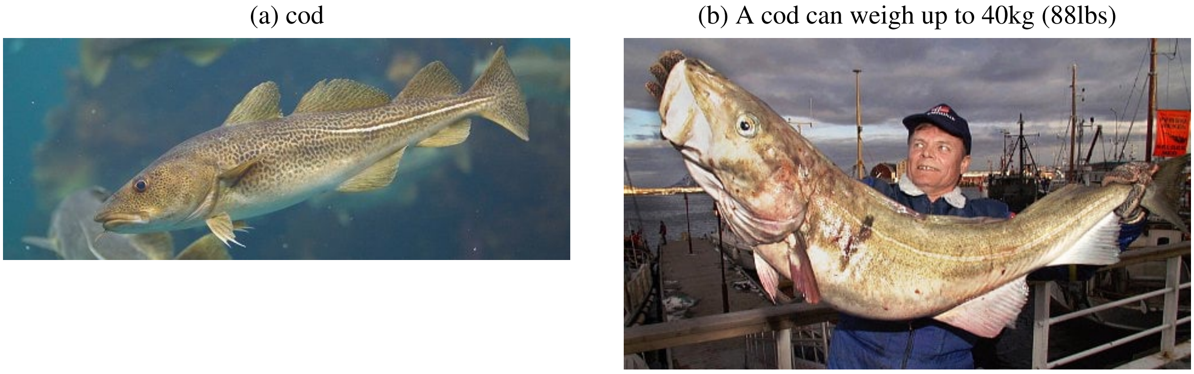
Tietenberg, Tom. 2003. “The tradable-permits approach to protecting the commons: Lessons for climate change.” *Oxford Review of Economic Policy*, 19(3): 400–419.

Appendices

A Atlantic cod

Figure A1 shows how cod looks like. Cod in Norwegian sea is Atlantic cod, scientific name *Gadus morhua*. They can live for 25 years and usually attain sexual maturity between two and four years old. They can grow to 1.3m and 40kg (88lbs). Atlantic cod is one of the most heavily fished species. It was fished for a thousand years by north European fishers who followed it across the North Atlantic Ocean to North America. It supported the US and Canada fishing economy until 1992, when fishing cod was limited. Several cod stocks collapsed in the 1990s (declined by more than 95% of maximum historical biomass) and have failed to fully recover even with the cessation of fishing.[17]

Figure A2 illustrates the distribution area and spawning area in Norwegian sea. The amount (numbers and biomass) increases from south to north, and around 75% lives north of the 62 latitude (the fishing areas that are studied in this paper).[18] The cod spawns in most of the fjords or in fjord arms in bigger fjord systems (within 200km from the coast).


(a) cod

(b) A cod can weigh up to 40kg (88lbs)

Figure A1: Cod

B Elasticities when firms are not price takers in input markets

Proposition 2 shows the relation between output elasticity of cost and input elasticity of output in a perfectly competitive input market. I now discuss the relation when input price depends on the input usage of the firms. In this case, the cost minimization problem in the

---

[17] See Frank et al. (2005) and NOAA, `https://www.fisheries.noaa.gov/species/atlantic-cod`

[18] See the description by the Institute of Marine Research, `https://www.hi.no/en/hi/temasider/species/costal-cod--north-of-the-62-latitude`.

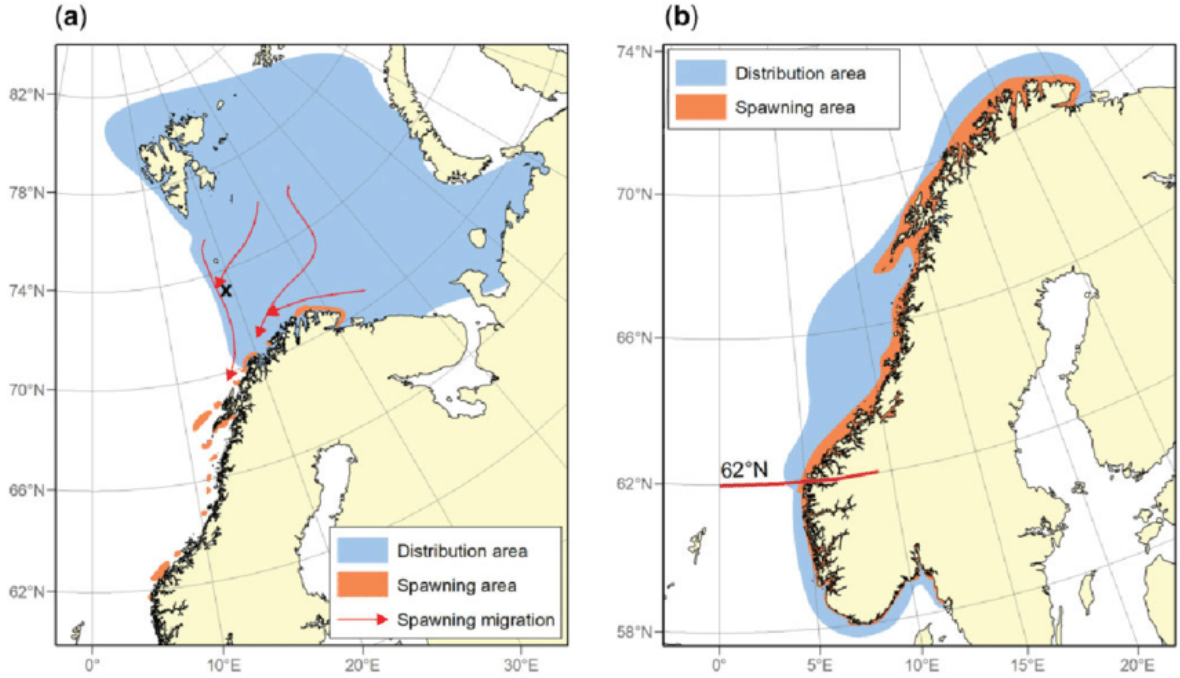


Figure A2: Cod fishery for the area north of 62°N

production stage will be:

$$C_{it}(q_{it}) = \min_{\boldsymbol{X}_{it}} \sum_{X \in \mathbb{X}} W_{it}^{X}(X_{it}) X_{it} \text{ subject to } q_{it} \leq Q_{it}(\boldsymbol{X}_{it}), \tag{32}$$

where $W_{it}^{X}(\cdot)$ is the input price function of the X-input use.[19] The Lagrangian function of the cost-minimization problem is

$$\mathcal{L}_{it} = \sum_{X \in \mathbb{X}} W_{it}^{X}(X_{it}) X_{it} + \lambda_{it}(q - Q_{it}(\boldsymbol{X}_{it})). \tag{33}$$

The first-order condition for any input $X \in \mathbb{X}$ is

$$\frac{\partial \mathcal{L}_{it}}{\partial X_{it}} = W_{it}^{X'} X_{it} + W_{it}^{X} - \lambda_{it} \cdot \frac{\partial Q_{it}}{\partial X_{it}} = 0. \tag{34}$$

---

[19] In this problem, I implicitly assume each input market is independent of each other. If the input markets are interdependent, the cost minimization will be

$$C_{it}(q_{it}) = \min_{\boldsymbol{X}_{it}} \boldsymbol{W}_{it}^{\top}(\boldsymbol{X}_{it}) \cdot \boldsymbol{X}_{it} \text{ subject to } q_{it} \leq Q_{it}(\boldsymbol{X}_{it}),$$

where $\boldsymbol{X}_{it}$ is the input vector and $\boldsymbol{W}_{it}(\cdot)$ is the vector form of the input-price function that depends on all types of input.

Summing this relation for all inputs and making a few algebra transformation, we arrive in:

$$\phi_{it} = \left(1 + \sum_{X \in \mathbb{X}} \eta_X \cdot \frac{W^X_{it} X_{it}}{C_{it}}\right) \left(\sum_{X \in \mathbb{X}} \theta^X_{it}\right)^{-1}, \tag{35}$$

where η_X is the price elasticity of demand for input X, i.e. $\eta_X \equiv \frac{dW^X}{dX} \cdot \frac{X}{W^X}$.

So, the output elasticity of cost is the ratio of average price elasticity of input demand, weighted by the share of input cost in total cost, to total input elasticities of output.

C Cost minimization and profit maximization

Proof of proposition 4. First, consider the Cournot competition. The profit maximization problem is

$$\max_{\boldsymbol{X}_{it}} P(\mathcal{Q}(Q_{it}(\boldsymbol{X}_{it}))) \cdot Q_{it}(\boldsymbol{X}_{it}) - G(\boldsymbol{X}_{it}),$$

where $Q_{it}(\boldsymbol{X}_{it})$ is the production function that defines the output quantity the firm can produce with such input use. The profit equals the revenue, which is the product of market price and the firm's output, subtracted by the cost $G(\cdot)$ the firm pays for their input uses. In the Cournot market environment, the market price depends on the total output of all firms in the market $\mathcal{Q}$. Of course, we have $\frac{d\mathcal{Q}}{dQ_{it}} = 1$, because $\mathcal{Q}$ is the industry output. Assume differentiability for all functions and concavity of the profit function, the optimal input use to maximize profits satisfies the following first order condition:

$$P' \cdot \frac{\partial Q}{\partial X} \cdot Q + P \cdot \frac{\partial Q}{\partial X} - \frac{\partial G}{\partial X} = 0.$$

Now, consider the alternative two-step decision process. In the first stage, the firm decides the output level that maximizes the following profits:

$$\max_{q_{it}} P(\mathcal{Q}(q_{it})) \cdot q_{it} - C(q_{it}),$$

where $C(q_{it})$ is the cost of producing q_{it} units of output. In the second stage, the firm decides the input use to minimize this cost of producing q_{it}. That is,

$$\min_{\boldsymbol{X}_{it}} G(\boldsymbol{X}_{it}) \text{ subject to } Q_{it}(\boldsymbol{X}_{it}) \geq q_{it}.$$

The optimal output and input levels in the two-step decision process satisfy the following first order conditions:

$$\begin{aligned}
P' \cdot q + P - C' &= 0, \\
\frac{\partial G}{\partial X_{it}} - \lambda \frac{\partial Q_{it}}{\partial X_{it}} &= 0, \\
Q_{it}(\boldsymbol{X}_{it}) &= q_{it} \text{ (assumming interior solutions),}
\end{aligned}$$

where λ is the multiplier associated with the targeted output constraint. Notice that the marginal cost C' is the shadow price of output constraint λ. The three conditions imply $P' \cdot Q + P - \frac{\partial G / \partial X}{\partial Q / \partial X} = 0$, which is equivalent to the first order condition of the profit-maximizing input-choice problem. Hence, the two decision problems, input choice to maximize profits and 2-step decision to maximize profits and minimize production cost, are equivalent in the Cournot market environment.

Now, consider the case where price is endogenous in output due to bargaining power. The profit maximization problem is

$$\max_{\boldsymbol{X}_{it}} P(\mathcal{Q}(Q_{it}(\boldsymbol{X}_{it})), Q_{it}(\boldsymbol{X}_{it})) \cdot Q_{it}(\boldsymbol{X}_{it}) - G(\boldsymbol{X}_{it}).$$

The profit-maximizing input must satisfy

$$\left(P_1 \cdot \frac{\partial Q}{\partial X} + P_2 \cdot \frac{\partial Q}{\partial X} \right) \cdot Q + P \cdot \frac{\partial Q}{\partial X} - \frac{\partial G}{\partial X} = 0,$$

where P_1, P_2 denote partial derivatives: $P_1 = \frac{\partial P}{\partial \mathcal{Q}}, P_2 = \frac{\partial P}{\partial Q}$.

Consider the two-step decision

$$\begin{aligned}
\max_{q_{it}} P(\mathcal{Q}(q_{it}), q_{it}) \cdot q_{it} - C(q_{it}) &\text{ in stage 1, and} \\
C(q_{it}) = \min_{\boldsymbol{X}_{it}} G(\boldsymbol{X}_{it}) \text{ subject to } Q_{it}(\boldsymbol{X}_{it}) \geq q_{it} &\text{ in stage 2.}
\end{aligned}$$

The optimal output and input in the two-step decision must satisfy

$$\begin{aligned}
(P_1 + P_2) \cdot q + P - \frac{dC}{dq} &= 0, \\
\frac{\partial G}{\partial X} - \lambda \frac{\partial Q}{\partial X} &= 0, \\
Q(\boldsymbol{X}) &= q.
\end{aligned}$$

Because the marginal cost is the shadow price $\frac{dC}{dq} = \lambda$, the three above conditions imply the first-order-condition of the profit-maximization problem. So, the two decision problems are equivalent in the presence of bargaining power.

Consider the third situation in which price is endogenous in product quality H and the quality can be adjusted by effort e_{it}. Then the equivalent two-step decision is

$$\max_{q_{it},e_{it}} P(\mathcal{Q}(q_{it}), q_{it}, H(e_{it})) \cdot q_{it} - C(q_{it}, e_{it}) \text{ in stage 1, and}$$
$$C(q_{it}, e_{it}) = \min_{\boldsymbol{X}_{it}} G(\boldsymbol{X}_{it}, e_{it}) \text{ subject to } Q_{it}(\boldsymbol{X}_{it}) \geq q_{it} \text{ in stage 2.}$$

The reason is the optimal output, effort, and input must satisfy

$$\begin{aligned}(P_1 + P_2) \cdot q + P - C_1 &= 0,\\ P_3 \cdot q - C_2 &= 0,\\ \frac{\partial G}{\partial X} - \lambda \frac{\partial Q}{\partial X} &= 0.\end{aligned}$$

Because $\frac{\partial C}{\partial q} = \lambda$ and $\frac{\partial C}{\partial e} = \frac{\partial G}{\partial e}$, the three above conditions imply the two first-order conditions that input and effort in the profit-maximization problem satisfy.

However, in a price-differentiation environment where the firm can use its production input to adjust product quality, the two problems, profit-maximizing input choice and two-step decision, are not equivalent in general. That is, consider the case $P_{it} = P(\mathcal{Q}(Q_{it}), Q_{it}, H(\boldsymbol{X}_{it}))$, where product quality $H(\cdot)$ can be directly adjusted by the production input factors $\boldsymbol{X}_{it}$. In this environment, there does not exist an equivalent two-step decision with the cost-minimizing input choice in the second stage, unless the quality function $H(\cdot)$ satisfies a set of conditions in relation to the price function and the production function $Q(\cdot)$.

Finally, consider the flexible form of the cost function in which output and input are interdependent. In this environment, the profit-maximizing input-choice problem is

$$\max_{\boldsymbol{X}_{it}} P(\mathcal{Q}(Q_{it}(\boldsymbol{X}_{it})), Q_{it}(\boldsymbol{X}_{it})) \cdot Q_{it}(\boldsymbol{X}_{it}) - G(Q_{it}(\boldsymbol{X}_{it}), \boldsymbol{X}_{it}).$$

The input choice must satisfy

$$\left(P_1 \cdot \frac{\partial Q}{\partial X} + P_2 \cdot \frac{\partial Q}{\partial X}\right) \cdot Q - P \cdot \frac{\partial Q}{\partial X} - \frac{\partial G}{\partial Q} \cdot \frac{\partial Q}{\partial X} - \frac{\partial G}{\partial X} = 0.$$

The equivalent two-step decision is

$$\max_{q_{it}} P(\mathcal{Q}(q_{it}), q_{it}) \cdot q_{it} - C(q_{it}) \text{ in stage 1, and}$$
$$C(q_{it}) = \min_{\boldsymbol{X}_{it}} G(q_{it}, \boldsymbol{X}_{it}) \text{ subject to } Q_{it}(\boldsymbol{X}_{it}) \geq q_{it} \text{ in stage 2,}$$

where the output and input must satisfy

$$\begin{aligned}(P_1+P_2)\cdot q+P-C'&=0,\\ \frac{\partial G}{\partial X}-\lambda\cdot\frac{\partial Q}{\partial X}&=0,\\ Q(\boldsymbol{X})&=q.\end{aligned}$$

Because $\frac{dC}{dq}=\frac{\partial G}{\partial q}+\lambda$, these three conditions imply the first-order condition of the profit maximizing problem. Hence, the two problems are equivalent.

D Decomposition Method

Proof of proposition 6.

Denote the production function $Q_{it}=\mathbb{F}(\mathbf{X}_{it},\omega_{it})$.[20] The change in output between $t-1$ and t is

$$\begin{aligned}\Delta Q_{it}&\equiv Q_{it}-Q_{i,t-1} &(36)\\ &=\mathbb{F}(\mathbf{X}_{it},\omega_{it})-\mathbb{F}(\mathbf{X}_{i,t-1},\omega_{i,t-1}) &(37)\\ &=\underbrace{\mathbb{F}(\mathbf{X}_{it},\omega_{it})-\mathbb{F}(\mathbf{X}_{i,t-1},\omega_{it})}_{\text{input reallocation }\Delta Q|_{\Delta X}}+\underbrace{\mathbb{F}(\mathbf{X}_{i,t-1},\omega_{it})-\mathbb{F}(\mathbf{X}_{i,t-1},\omega_{i,t-1})}_{\text{productivity change }\Delta Q|_{\Delta\omega}}. &(38)\end{aligned}$$

The first term "input reallocation" in this decomposition represents the change in output due to input adjustment given the new realized productivity. The second term "productivity change" represents the change due to change in productivity between two periods, given no input adjustment.

I now show that first term captures the slide on the average cost curve, whereas the second term captures the shift in the average costs. Notice that the production cost $C(q)$ is the cost of inputs evaluated at cost-minimized input choices $X^*(q,\omega,\mathbf{W},\beta)$. Hence, every point $\tilde{X}^*(\tilde{q},\omega,\mathbf{W},\beta)$ gives a unique point $(\tilde{q},\tilde{AC})$ on an average cost curve and vice versa. If the production function contains only Hicksian productivity, then output elasticity of input $\theta_X=\frac{\partial\ln Q}{\partial\ln X}$ does not depend on productivity ω, resulting cost elasticity ϕ being independent of ω. Hence, the same input choices, regardless of productivity levels, generate the same cost elasticity.

---

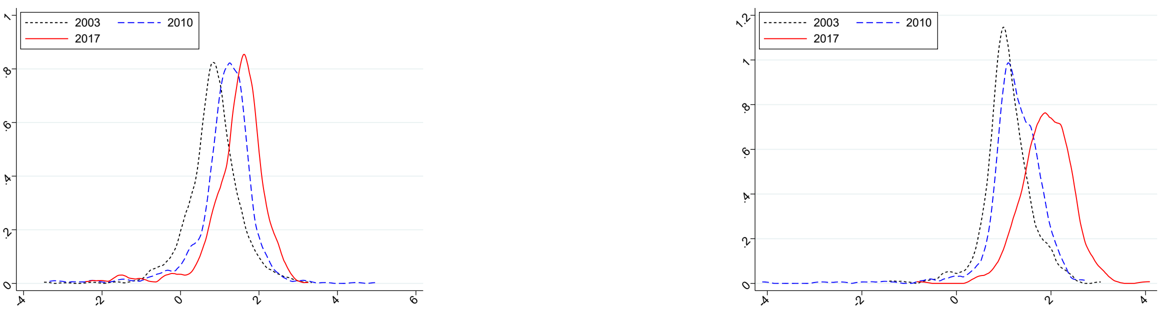
[20] The output growth is $\ln Q_{it}=\mathcal{F}(\mathbf{X}_t,\omega_t)=\ln\mathbb{F}(\cdot)$. The decomposition of output growth follows the same method as the decomposition of the output.

E Productivity and Cost Elasticity Using OLS with FEs and Dynamic Panel Approaches

Table E1: Summary statistics of estimates of cost indices by licensed length group

	Pre-trade-program					Post-trade-program				
	count	mean	sd	min	max	count	mean	sd	min	max
Panel A: Output elasticity of total costs, using the dynamic panel estimator for the production function										
0–10.9m	8,470	0.368	0.046	0.205	0.584	8,362	0.373	0.056	0.205	0.806
11–14.9m	3,590	0.429	0.058	0.249	0.660	3,637	0.445	0.071	0.225	0.790
15–20.9m	995	0.458	0.063	0.254	0.688	2,367	0.513	0.086	0.281	0.915
21–27.9m	433	0.463	0.066	0.300	0.734	1,016	0.532	0.095	0.326	1.005
Panel B: Output elasticity of total costs, using the OLS-FE estimator for the production function										
0–10.9m	8,470	0.387	0.054	0.205	0.651	8,362	0.393	0.066	0.205	0.969
11–14.9m	3590	0.456	0.070	0.249	0.741	3,637	0.476	0.085	0.224	0.934
15–20.9m	995	0.487	0.077	0.252	0.793	2,367	0.554	0.107	0.284	1.075
21–27.9m	433	0.485	0.079	0.298	0.824	1,016	0.567	0.119	0.323	1.217

Note: Pre-trade-program period and post-trade period for licensed groups 15–20.9m and 21–27.9m are 2001–2003 and 2005–2017. For licensed length group 11–14.9m, they are 2001–2007 and 2009–2017. Licensed length group 0–10.9m is not allowed to trade during 2001 and 2017, but we compare period 2001–2007 to period 2009–2017.


(a) productivity in licensed length 0–10.9m

(b) productivity in licensed length 11–14.9m


(c) productivity in licensed length 15–20.9m

(d) productivity in licensed length 21–27.9m

Figure E3: Distribution of productivity (from the OLS-with-FEs estimator)

(a) productivity in licensed length 0–10.9m (b) productivity in licensed length 11–14.9m

(c) productivity in licensed length 15–20.9m (d) productivity in licensed length 21–27.9m

Figure E4: Distribution of productivity (from the dynamic panel estimator)

(a) elasticity of total costs, license 0–10.9m (b) elasticity of total costs, license 11–14.9m

(c) elasticity of total costs, license 15–20.9m (d) elasticity of total costs, license 21–27.9m

Figure E5: Distribution of output elasticity of total costs (implied from the OLS with FEs)

(a) elasticity of total costs, license 0–10.9m

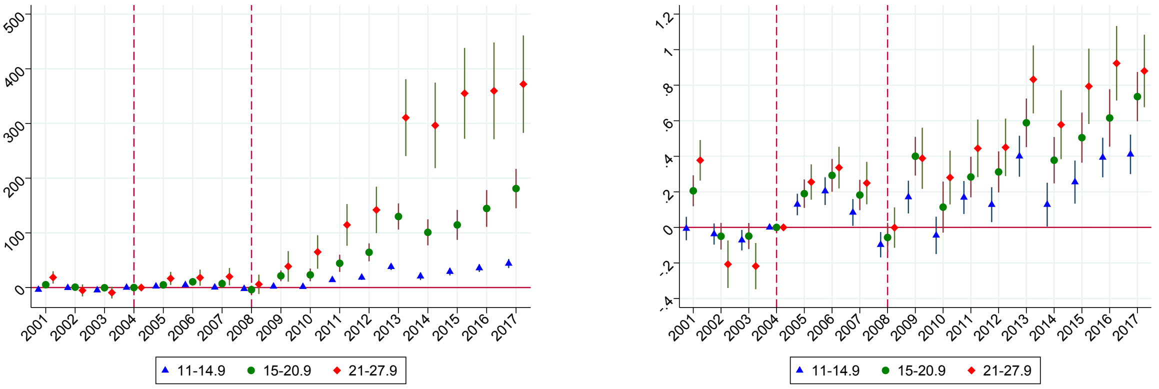
(b) elasticity of total costs, license 11–14.9m

(c) elasticity of total costs, license 15–20.9m

(d) elasticity of total costs, license 21–27.9m

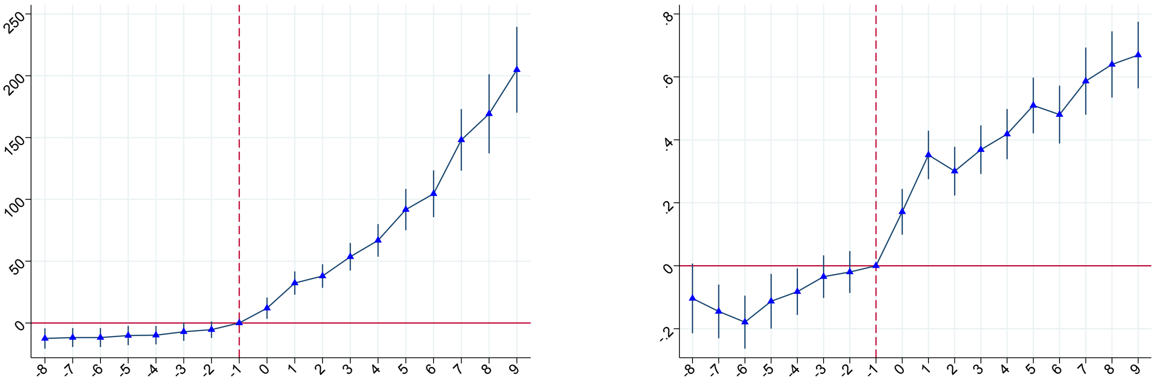
Figure E6: Distribution of output elasticity of total costs (implied from the OLS with FEs)

F Supplementary Event Studies and Diff-in-diff Results



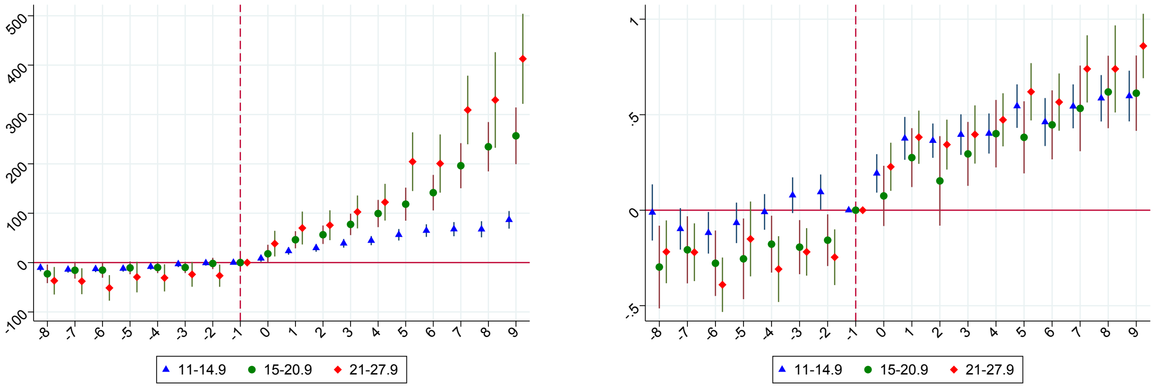
(a) catch quantity (tonne)

(b) logged quantity (tonne)



(c) catch quantity (tonne)

(d) logged quantity (tonne)

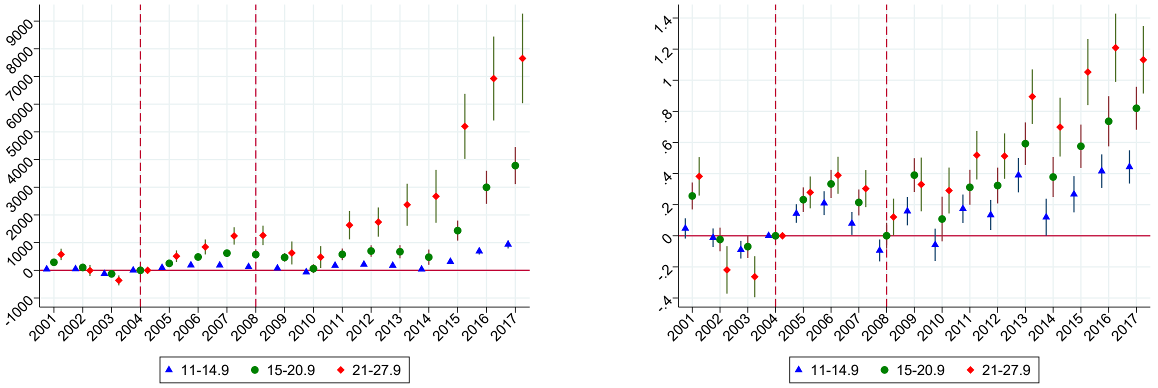


(e) catch quantity (tonne)

(f) logged quantity (tonne)

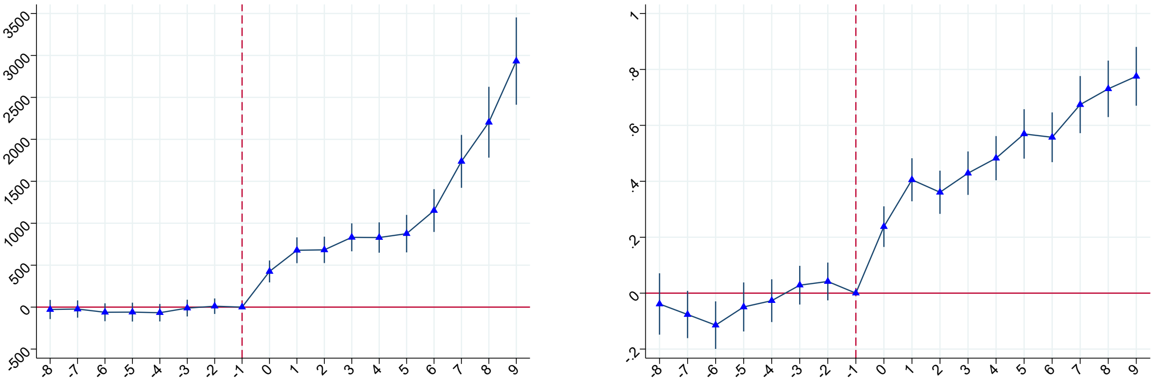
Figure F7: Impacts of the trading policy and quota acquisition on catch quantity and revenue

Note: Panels A and B plot the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panels C–F plot the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.



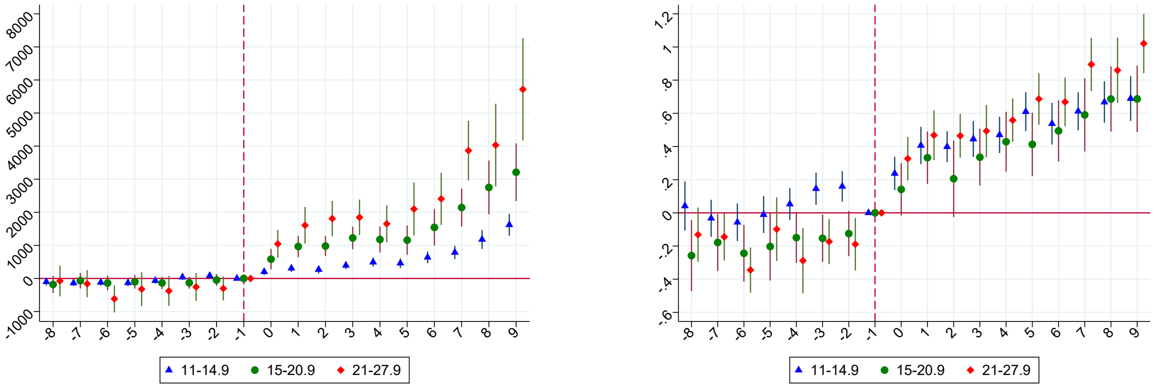
(a) revenue (thousand NOK)

(b) logged revenue



(c) revenue (thousand NOK)

(d) logged revenue

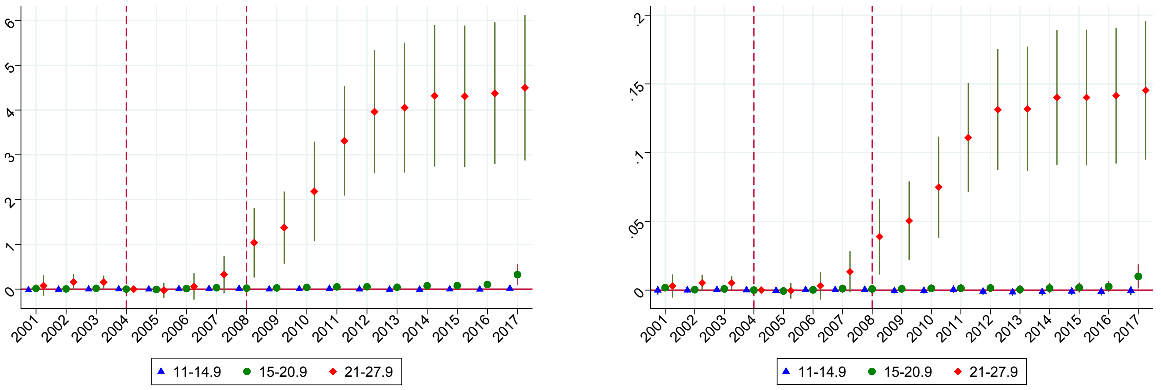


(e) revenue (thousand NOK)

(f) logged revenue

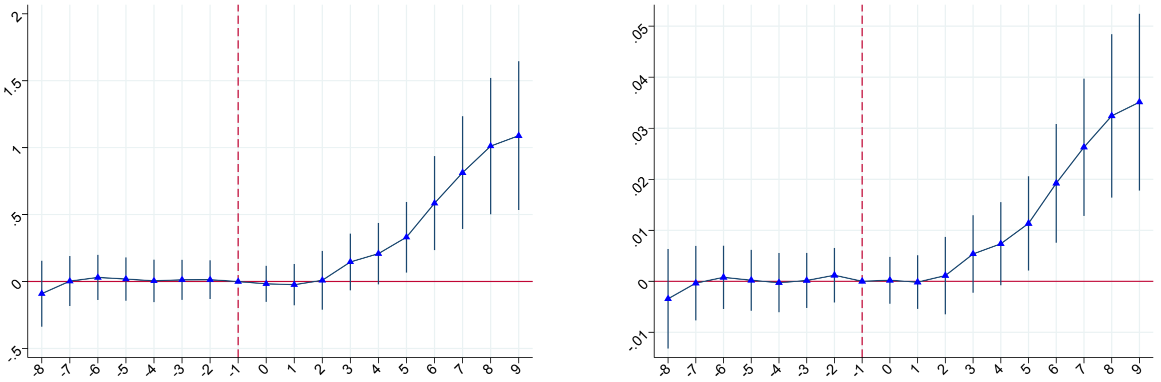
Figure F8: Impacts of the trading policy and quota acquisition on revenue

Note: Panels A and B plot the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panels C–F plot the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.



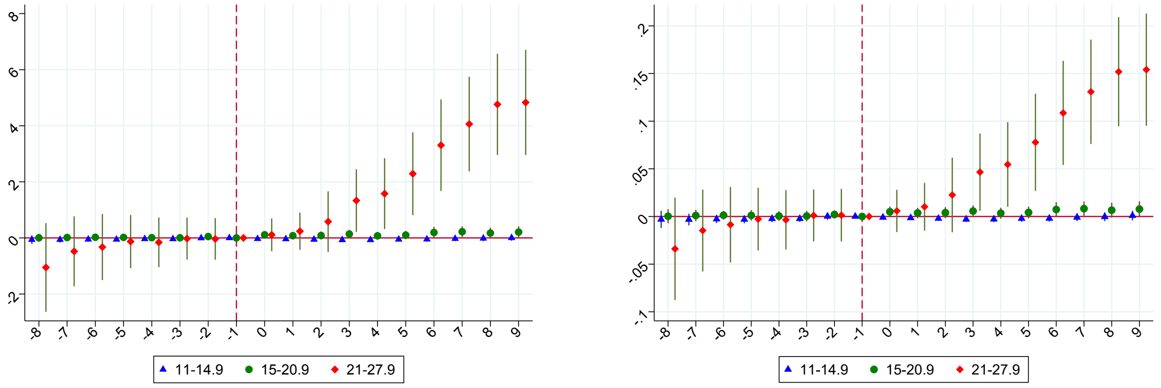
(a) length (m)

(b) logged length



(c) length (m)

(d) logged length

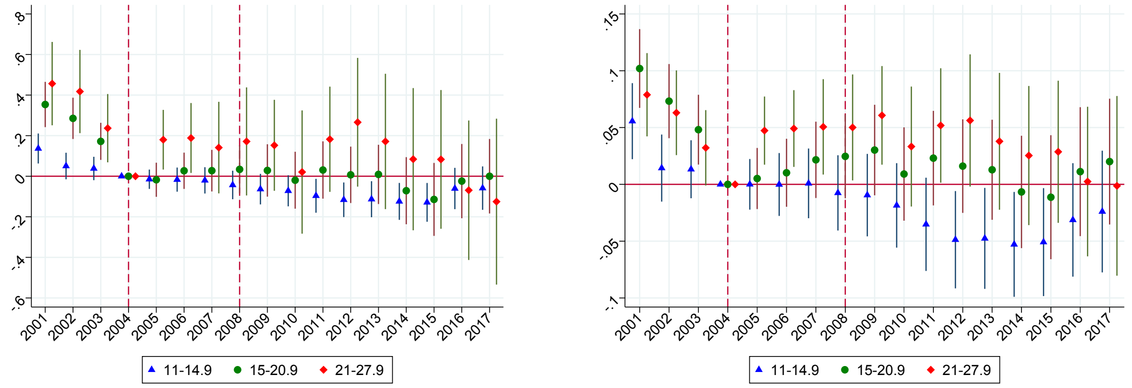


(e) length (m)

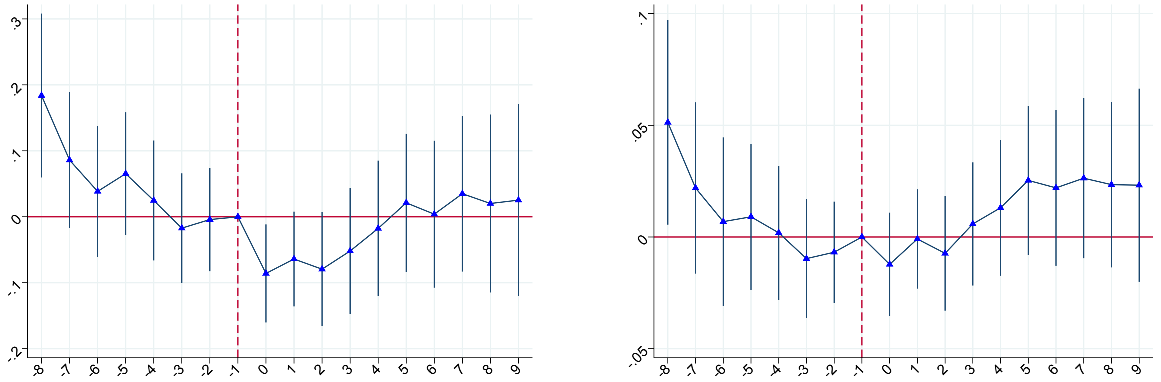
(f) logged length

Figure F9: Impacts of the trading policy and quota acquisition on vessel actual length

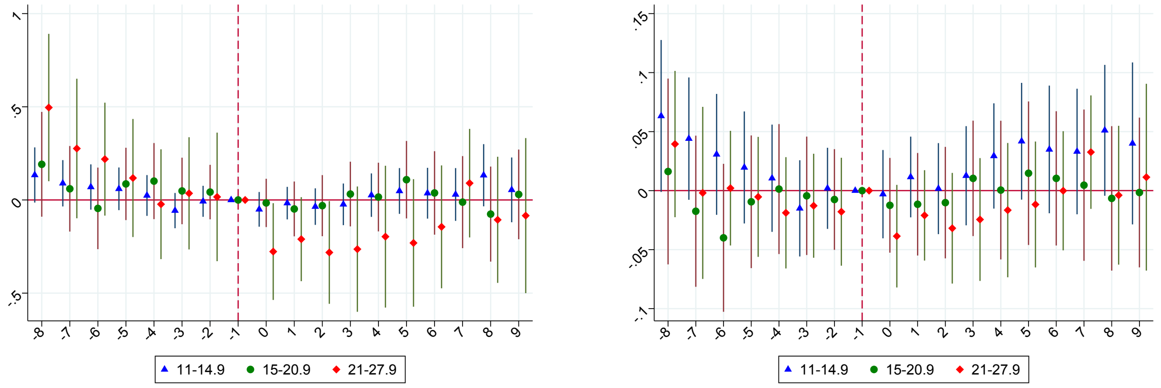
Note: Panels A and B plot the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panels C–F plot the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.


(a) crew (person)

(b) logged crew


(c) crew (person)

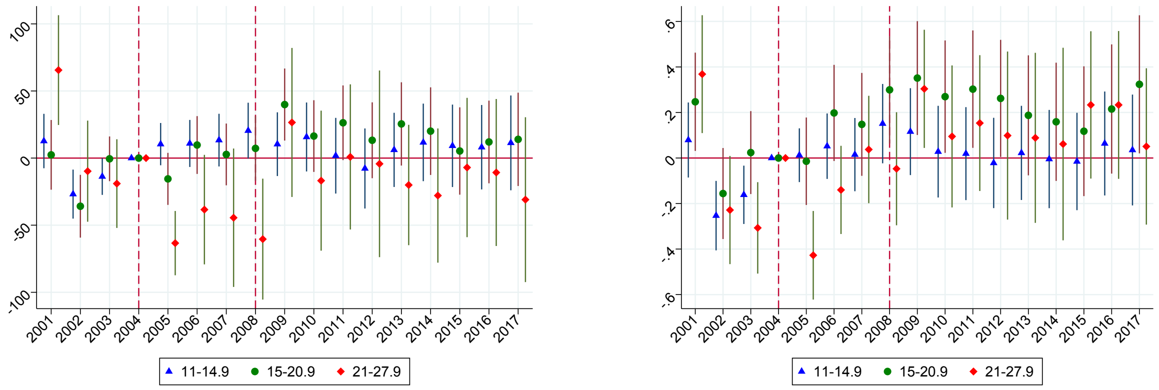
(d) logged crew


(e) crew (person)

(f) logged crew

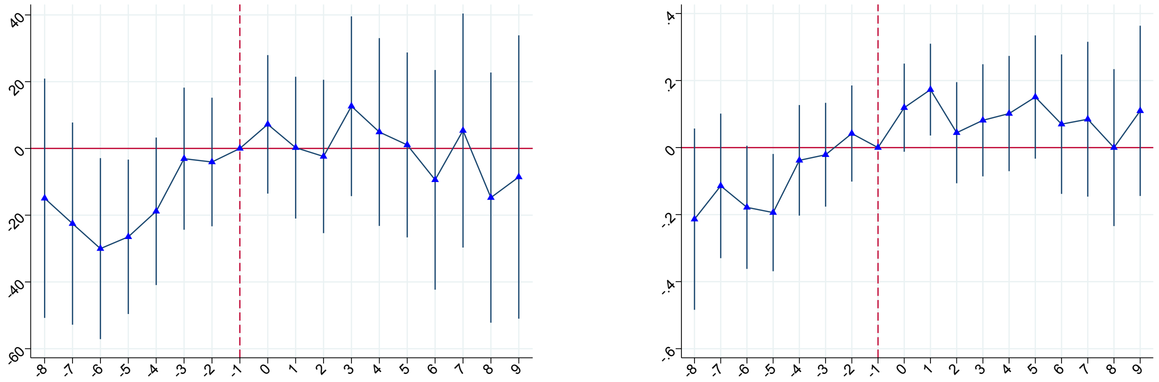
Figure F10: Impacts of the trading policy and quota acquisition on crew size

Note: Panels A and B plot the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panels C–F plot the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.



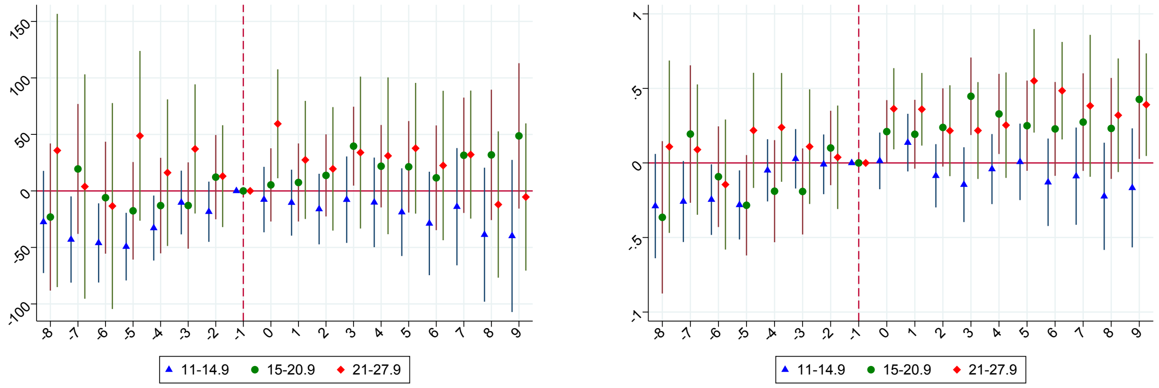
(a) distance (km)

(b) logged distance



(c) distance (km)

(d) logged distance

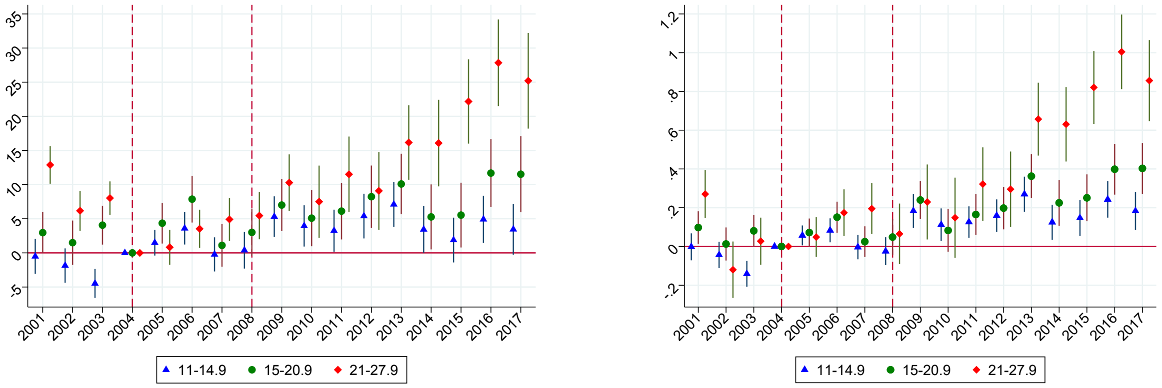


(e) distance (km)

(f) logged distance

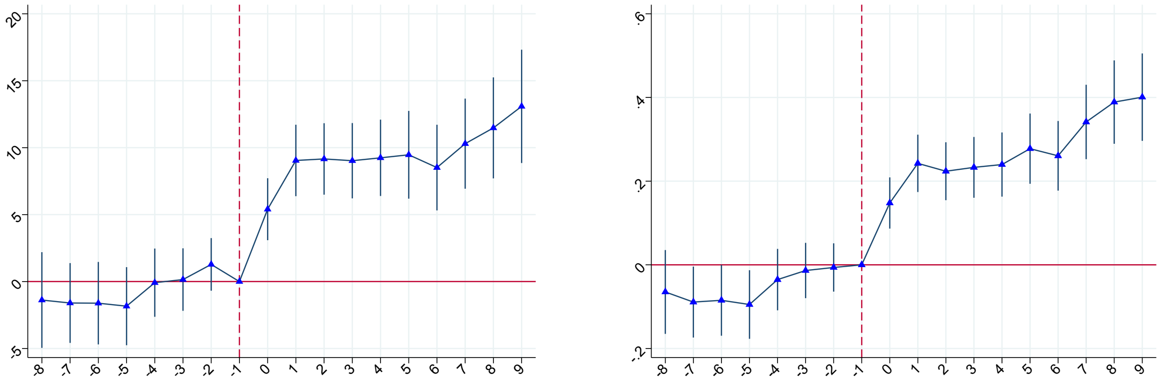
Figure F11: Impacts of the trading policy and quota acquisition on distance from fishers' municipality to major catch location

Note: Panels A and B plot the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panels C–F plot the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.



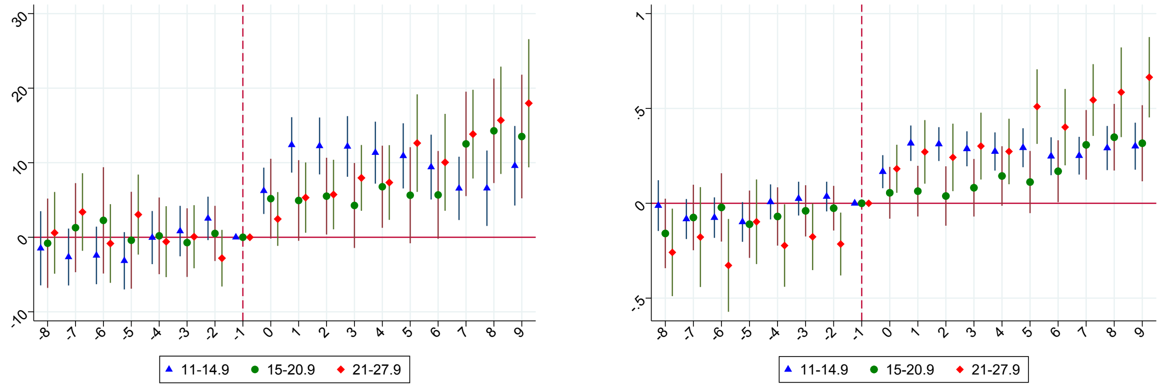
(a) # trips

(b) logged # trips



(c) # trips

(d) logged # trips

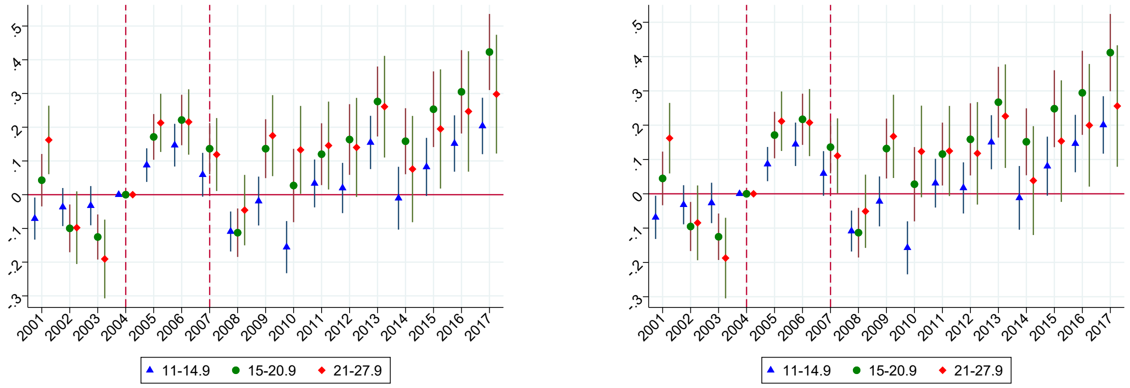


(e) # trips

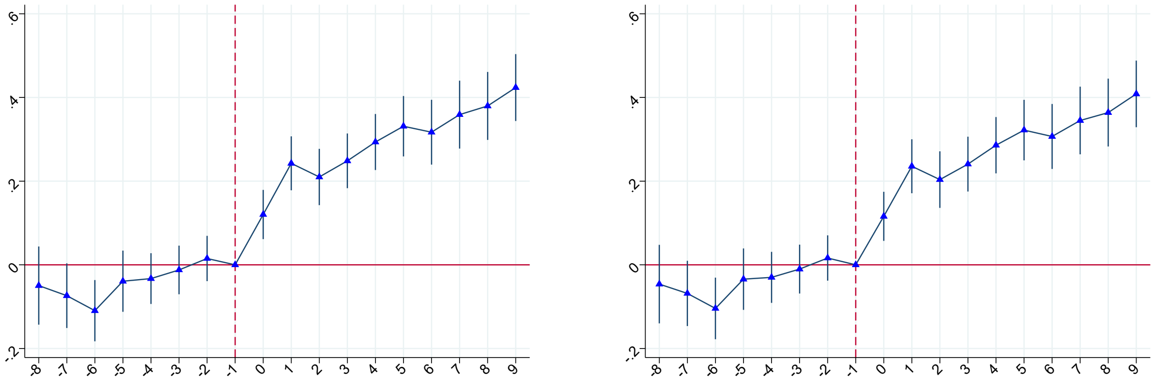
(f) logged # trips

Figure F12: Impacts of the trading policy and quota acquisition on the number of trips in a year

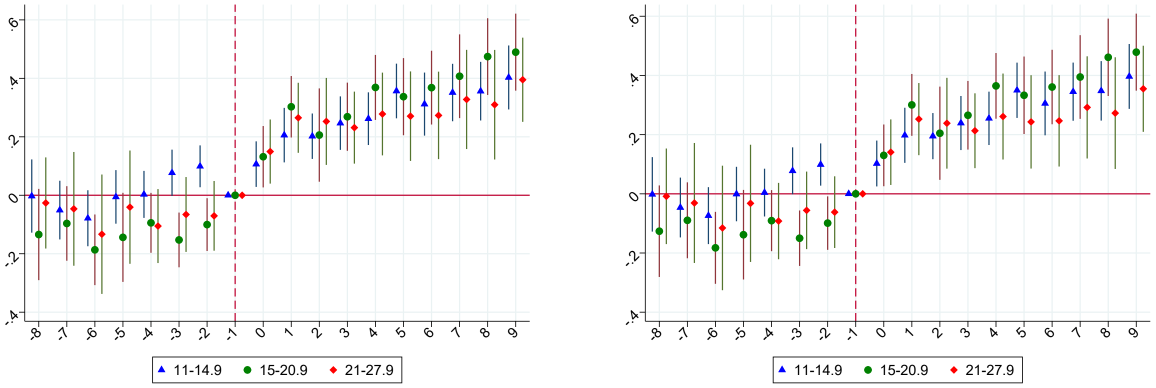
Note: Panels A and B plot the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panels C–F plot the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.


(a) productivity using OLS-with-FE estimator

(b) productivity using the proxy variable approach


(c) productivity using OLS with FEs

(d) productivity using the proxy variable approach

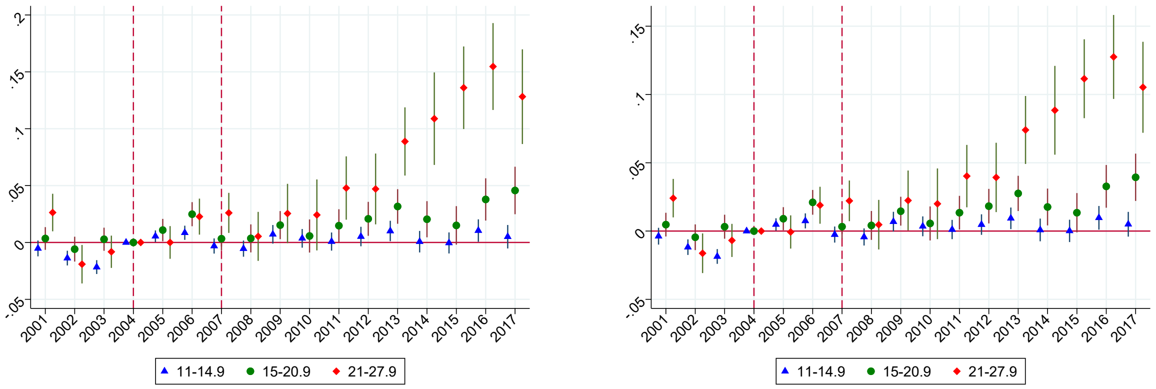

(e) productivity using OLS with FEs

(f) productivity using the proxy variable approach

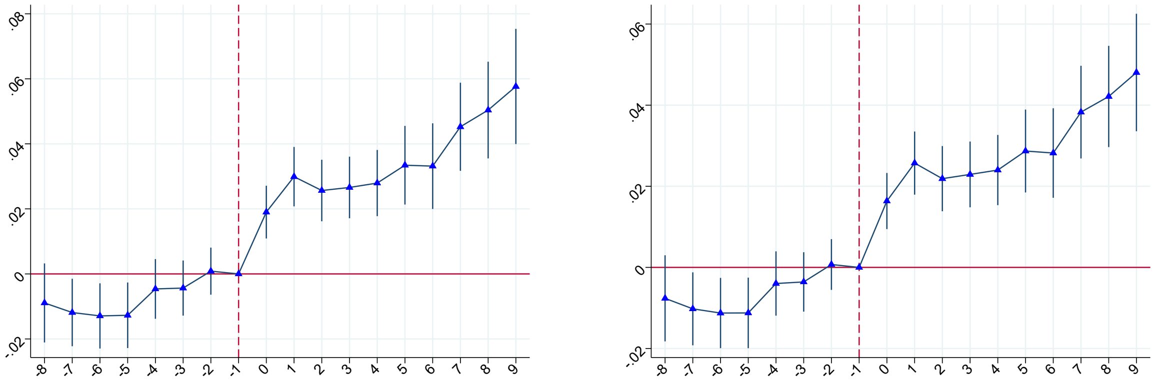
Figure F13: Impacts of the trading policy and quota acquisition on productivity

Note: Panels A and B plot the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panels C–F plot the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.

(a) ϕ implied from OLS with FEs (b) ϕ implied from the proxy variable approach



(c) ϕ implied from OLS with FEs (d) ϕ implied from the proxy variable approach



(e) ϕ implied from OLS with FEs (f) ϕ implied from the proxy variable approach

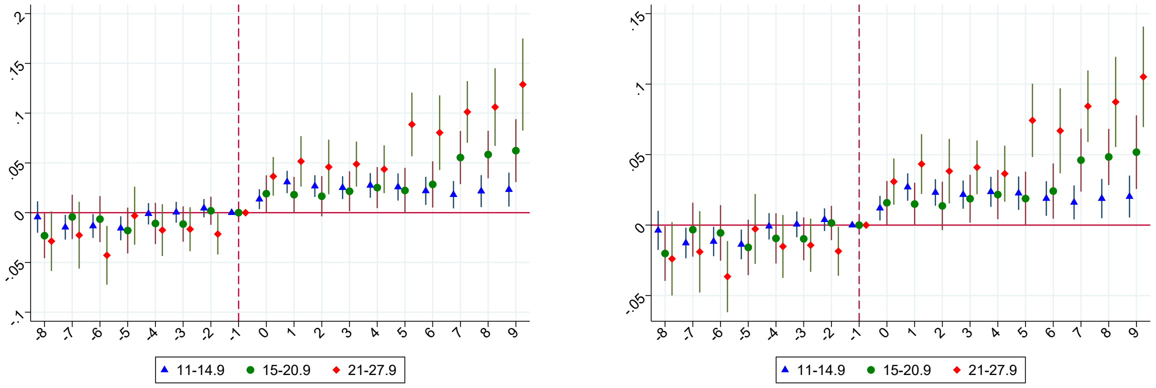


Figure F14: Impacts of the trading policy and quota acquisition on output elasticity of total cost

Note: Panels A and B plot the event study coefficients of ITT of the trading policy. Year 2004 and the non-tradable group (0–10.9m) are normalized. From 2004, the groups 15–20.9m and 21–27.9m may trade. From 2007, the group 11–14.9m may trade. Panels C–F plot the event study coefficients of ATT of quota acquisition, for years before and after the acquisition. Vessels in the non-tradable group and that are in the tradable group and do not trade are the base group.

Table F2: Effects of trading policy on production factors (log levels)

	(1) logged length (m)	(2) logged crew (person)	(3) logged distance (km)	(4) logged # trips
Panel A: ITT (pooling all trade qualified groups)				
Trade qualified	-0.001	-0.040***	0.039	0.072***
	(0.001)	(0.009)	(0.061)	(0.021)
Panel B: ITT by trade qualified group (license group)				
21-27.9m × From 2004	0.022***	-0.015	-0.029	0.060
	(0.007)	(0.014)	(0.084)	(0.064)
15-20.9m × From 2004	-0.001	-0.055***	0.059	0.009
	(0.001)	(0.015)	(0.100)	(0.038)
11-14.9m × From 2008	-0.005***	-0.038***	0.043	0.099***
	(0.002)	(0.011)	(0.071)	(0.024)
Panel C: pooled ATT using DID FE				
Quota acquisition	0.005*	0.008	0.169***	0.311***
	(0.003)	(0.009)	(0.065)	(0.027)
Panel D: ATT by license group, using DID FE				
Quota acquisition × 21-27.9m	0.042**	0.002	0.283**	0.463***
	(0.017)	(0.018)	(0.113)	(0.062)
Quota acquisition × 15-20.9m	0.001	0.004	0.287**	0.236***
	(0.002)	(0.019)	(0.124)	(0.044)
Quota acquisition × 11-14.9m	-0.003	0.011	0.086	0.304***
	(0.002)	(0.015)	(0.079)	(0.037)
Panel E: pooled LATE using IV DID FE				
Quota acquisition	-0.005	-0.184***	0.182	0.335***
	(0.005)	(0.046)	(0.284)	(0.094)
Kleibergen-Paap rk Wald F	131.991	131.991	131.991	131.991
Panel F: LATE by license group, using IV DID FE				
Quota acquisition × 21-27.9m	0.086***	-0.072	-0.102	0.242
	(0.027)	(0.053)	(0.339)	(0.244)
Quota acquisition × 15-20.9m	-0.005	-0.344***	0.367	0.070
	(0.006)	(0.116)	(0.589)	(0.234)
Quota acquisition × 11-14.9m	-0.014***	-0.175***	0.186	0.376***
	(0.005)	(0.050)	(0.289)	(0.090)
Kleibergen-Paap rk Wald F	13.406	13.406	13.406	13.406
Observations	30,776	30,776	30,776	30,776

Note: Panels represent specifications. Columns represent dependent variables. All specifications use yearly observations and include year fixed effects, vessel fixed effects, and owner fixed effects. Panels A and B estimate ITT of the trading policy using DID specifications in regression (23). Panels C and D estimate ATT of quota acquisition using DID with fixed effects specifications in regression (24). Panels E and F estimate ATT of quota acquisition using IV DID with fixed effects specification; the main regression is equation (24) but the treatment $\textit{Quota acquisition}_{it}$ is instrumented by the policy assignment $\textit{Trade qualified}_{it}$. Standard errors in parentheses are clustered by vessel's municipality. Significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F3: Effects of trading policy on productivity

	(1)	(2)	(3)
	OLS with FE	proxy variable	dynamic panel
	logged TFPQ ω	logged TFPQ ω	logged TFPQ ω
Panel A: ITT (pooling all trade qualified groups)			
Trade qualified	0.029	0.025	0.025
	(0.020)	(0.020)	(0.021)
Panel B: ITT by trade qualified group (license group)			
21-27.9m × From 2004	0.140^{***}	0.129^{***}	0.083^{*}
	(0.041)	(0.041)	(0.044)
15-20.9m × From 2004	0.147^{***}	0.145^{***}	0.150^{***}
	(0.032)	(0.033)	(0.034)
11-14.9m × From 2008	-0.037	-0.039	-0.033
	(0.024)	(0.024)	(0.024)
Panel C: pooled ATT using DID FE			
Quota acquisition	0.287^{***}	0.276^{***}	0.257^{***}
	(0.022)	(0.022)	(0.022)
Panel D: ATT by license group, using DID FE			
Quota acquisition × 21-27.9m	0.317^{***}	0.292^{***}	0.243^{***}
	(0.047)	(0.048)	(0.052)
Quota acquisition × 15-20.9m	0.390^{***}	0.381^{***}	0.358^{***}
	(0.039)	(0.039)	(0.037)
Quota acquisition × 11-14.9m	0.233^{***}	0.225^{***}	0.216^{***}
	(0.028)	(0.028)	(0.027)
Panel E: pooled LATE using IV DID FE			
Quota acquisition	0.133	0.118	0.117
	(0.094)	(0.094)	(0.097)
Kleibergen-Paap rk Wald F	131.991	131.991	131.991
Panel F: LATE by license group, using IV DID FE			
Quota acquisition × 21-27.9m	0.573^{***}	0.528^{***}	0.351^{**}
	(0.168)	(0.171)	(0.174)
Quota acquisition × 15-20.9m	0.915^{***}	0.898^{***}	0.929^{***}
	(0.202)	(0.203)	(0.218)
Quota acquisition × 11-14.9m	-0.002	-0.013	-0.002
	(0.094)	(0.094)	(0.098)
Kleibergen-Paap rk Wald F	13.406	13.406	13.406
Observations	30,776	30,776	30,776

Note: Panels represent specifications. Columns represent dependent variables. All specifications use yearly observations and include year fixed effects, vessel fixed effects, and owner fixed effects. Panels A and B estimate ITT of the trading policy using DID specifications in regression (23). Panels C and D estimate ATT of quota acquisition using DID with fixed effects specifications in regression (24). Panels E and F estimate ATT of quota acquisition using IV DID with fixed effects specification; the main regression is equation (24) but the treatment *Quota acquisition*$_{it}$ is instrumented by the policy assignment *Trade qualified*$_{it}$. Standard errors in parentheses are clustered by vessel's municipality. Significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F4: Effects of trading policy on fish sale prices

	(1) price (NOK/kg) (trip-level)	(2) logged price (trip-level)	(3) average value (NOK/kg) (yearly)	(4) logged avg value (yearly)
Panel A: ITT (pooling all trade qualified groups)				
Trade qualified	0.077	0.005	-0.194	-0.005
	(0.060)	(0.004)	(0.132)	(0.006)
Panel B: ITT by trade qualified group (license group)				
21-27.9m × From 2004	0.316**	0.019**	0.271	0.020*
	(0.150)	(0.008)	(0.242)	(0.012)
15-20.9m × From 2004	0.427**	0.026**	-0.170	0.001
	(0.164)	(0.010)	(0.314)	(0.013)
11-14.9m × From 2008	-0.061	-0.003	-0.285	-0.012
	(0.065)	(0.004)	(0.201)	(0.009)
Panel C: pooled ATT using DID FE				
Quota acquisition	-0.092	-0.006	-0.228*	-0.014**
	(0.074)	(0.005)	(0.122)	(0.006)
Panel D: ATT by license group, using DID FE				
Quota acquisition × 21-27.9m	-0.198	-0.009	-0.192	-0.011
	(0.140)	(0.010)	(0.225)	(0.013)
Quota acquisition × 15-20.9m	0.077	0.006	-0.082	-0.005
	(0.136)	(0.009)	(0.211)	(0.013)
Quota acquisition × 11-14.9m	-0.156*	-0.011*	-0.304**	-0.019**
	(0.088)	(0.006)	(0.137)	(0.008)
Panel E: pooled LATE using IV DID FE				
Quota acquisition	0.320	0.021	-0.904	-0.024
	(0.259)	(0.017)	(0.625)	(0.026)
Kleibergen-Paap rk Wald F	118.702	118.702	131.991	131.991
Panel F: LATE by license group, using IV DID FE				
Quota acquisition × 21-27.9m	1.300*	0.079**	1.022	0.079
	(0.717)	(0.040)	(1.018)	(0.051)
Quota acquisition × 15-20.9m	2.382***	0.146***	-1.051	0.006
	(0.628)	(0.036)	(2.043)	(0.082)
Quota acquisition × 11-14.9m	0.142	0.010	-1.068	-0.037
	(0.240)	(0.016)	(0.689)	(0.030)
Kleibergen-Paap rk Wald F	10.232	10.232	13.406	13.406
Observations	1,158,487	1,158,487	30,067	30,067

Note: Panels represent specifications. Columns represent dependent variables. All specifications use yearly observations and include year fixed effects, vessel fixed effects, and owner fixed effects. Panels A and B estimate ITT of the trading policy using DID specifications in regression (23). Panels C and D estimate ATT of quota acquisition using DID with fixed effects specifications in regression (24). Panels E and F estimate ATT of quota acquisition using IV DID with fixed effects specification; the main regression is equation (24) but the treatment $\textit{Quota acquisition}_{it}$ is instrumented by the policy assignment $\textit{Trade qualified}_{it}$. Standard errors in parentheses are clustered by vessel's municipality. Significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G Productivity

Table G5: Decomposition of change in productivity

	(1)	(2)	(3)	(4)	(5)
	2001–2014	2001–2004	2004–2007	2008–2011	2011–2014
Panel A: Licensed 0–10.9m					
Total	95.83	-6.83	16.50	29.51	29.06
Within	71.39	-3.32	1.43	22.27	21.93
Net entry	6.26	-1.76	3.46	1.37	1.55
Between	-3.50	-2.53	4.45	0.98	2.19
Covariance	21.67	0.79	7.17	4.89	3.38
Re-allocation	18.18	-1.75	11.62	5.87	5.58
Panel B: Licensed 11–14.9m					
Total	92.60	-1.88	10.65	27.78	26.30
Within	73.82	-3.49	0.92	25.14	20.89
Net entry	6.31	2.49	2.88	-0.03	1.78
Between	3.76	0.24	2.57	0.22	2.24
Covariance	8.70	-1.12	4.28	2.45	1.39
Re-allocation	12.46	-0.88	6.85	2.67	3.63
Panel C: Licensed 15–20.9m					
Total	108.55	-3.75	16.20	30.26	20.67
Within	80.43	-3.96	3.71	24.85	18.31
Net entry	18.74	1.12	7.38	3.03	2.84
Between	3.83	1.23	2.97	0.74	-0.45
Covariance	5.55	-2.13	2.15	1.64	-0.04
Re-allocation	9.38	-0.91	5.12	2.37	-0.49
Panel D: Licensed 21–27.9m					
Total	88.90	4.87	9.33	21.12	17.67
Within	57.11	-2.12	2.39	17.45	13.74
Net entry	16.24	6.04	3.82	-0.28	2.68
Between	6.93	-0.48	2.29	1.27	0.92
Covariance	8.62	1.44	0.84	2.69	0.33
Re-allocation	15.55	0.96	3.12	3.95	1.26

Note: Selected periods, percentage changes during the period.

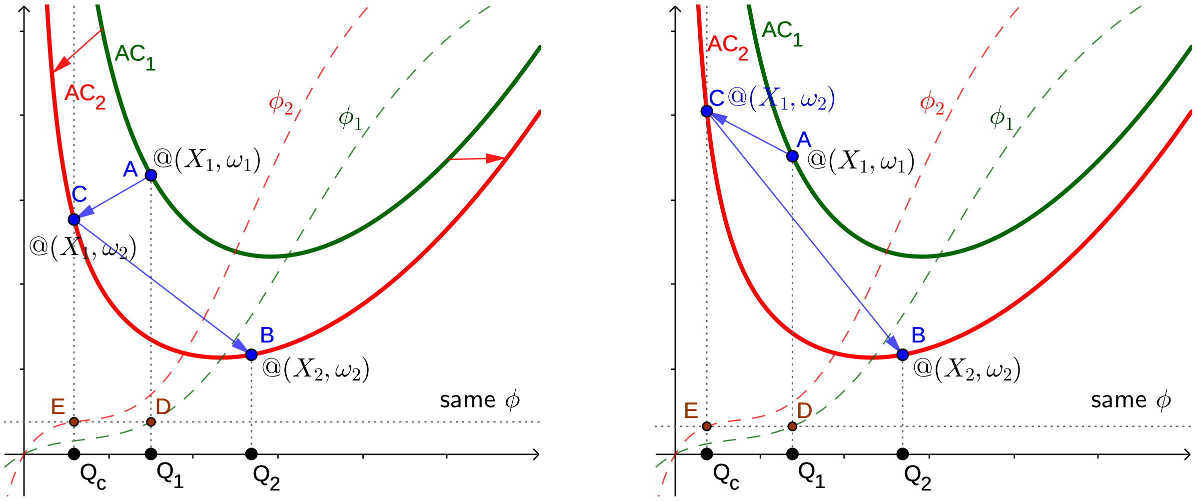
Table G6: Productivity index

	(1) Ever-trading continuer at t-k	(2) Ever-trading continuer at t	(3) Never-trading continuer at t-k	(4) Never-trading continuer at t	(5) Entrant at t	(6) Exiter at t-k
Licensed length 0–10.9m						
2001–2014			0.06	0.80	0.87	-0.06
2001–2004			-0.00	-0.05	-0.09	0.01
2004–2007			-0.05	0.06	0.15	-0.08
2008–2011			0.15	0.43	0.51	0.15
2011–2014			0.47	0.82	0.85	0.36
Licensed length 11–14.9m						
2001–2014	0.17	0.99	-0.03	0.60	1.00	-0.07
2001–2004			-0.01	-0.02	0.01	0.04
2004–2007			0.00	0.06	0.24	-0.12
2008–2011	0.33	0.65	0.16	0.42	0.63	0.19
2011–2014	0.67	1.02	0.47	0.78	1.10	0.47
Licensed length 15–20.9m						
2001–2014	0.16	1.12	-0.02	0.61	1.24	-0.02
2001–2004	0.32	0.28	0.01	-0.05	-0.04	-0.04
2004–2007	0.14	0.27	-0.07	-0.03	0.23	-0.10
2008–2011	0.50	0.87	0.27	0.61	0.85	0.27
2011–2014	0.83	1.20	0.69	0.96	1.52	0.75
Licensed length 21–27.9m						
2001–2014	0.09	0.95	0.05	0.45	1.09	-0.07
2001–2004	0.20	0.27	-0.00	-0.02	0.25	-0.04
2004–2007	0.21	0.28	0.03	0.07	0.14	-0.11
2008–2011	0.43	0.68	0.38	0.66	0.76	0.30
2011–2014	0.81	1.05	0.69	1.00	1.06	0.60

H Economies of scale vs. cost shifting

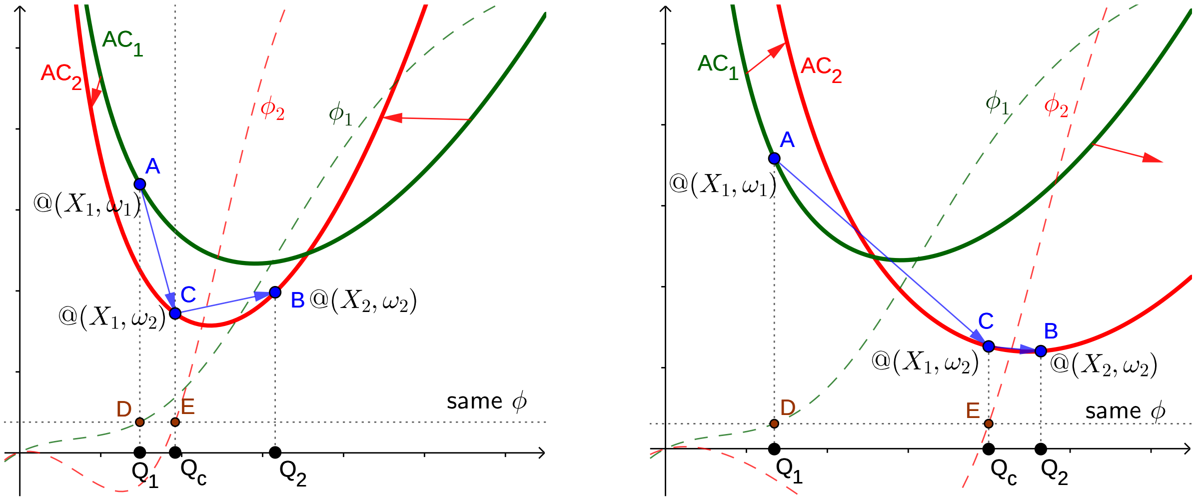
(a) average cost shifts inward and rightward

(b) average cost shifts inward and rightward



(c) average cost shifts inward and leftward

(d) average cost shifts outward and rightward



(e) average cost shifts outward and leftward

(f) average cost shifts outward and leftward

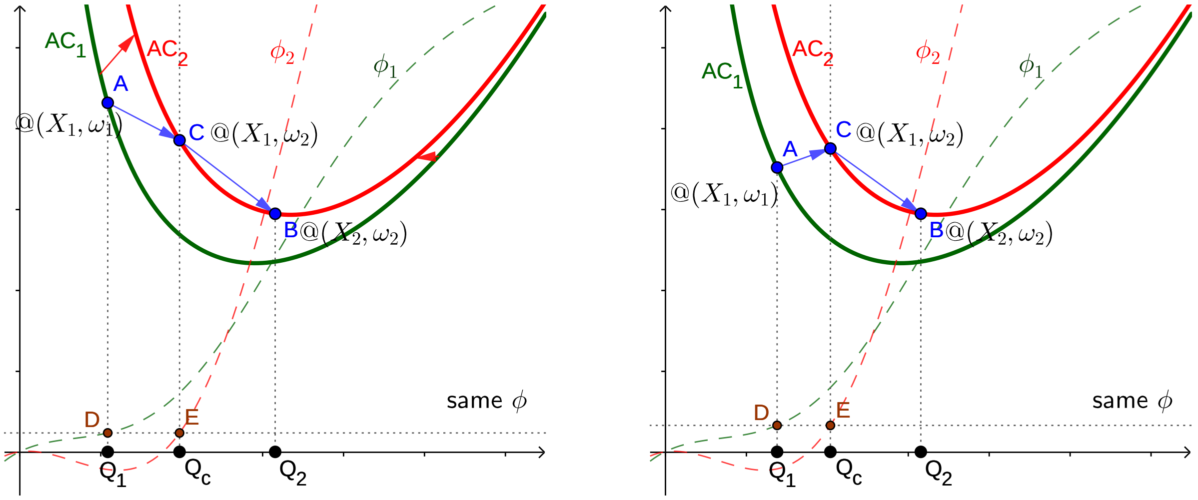


Figure H15: Decomposing gains from trading

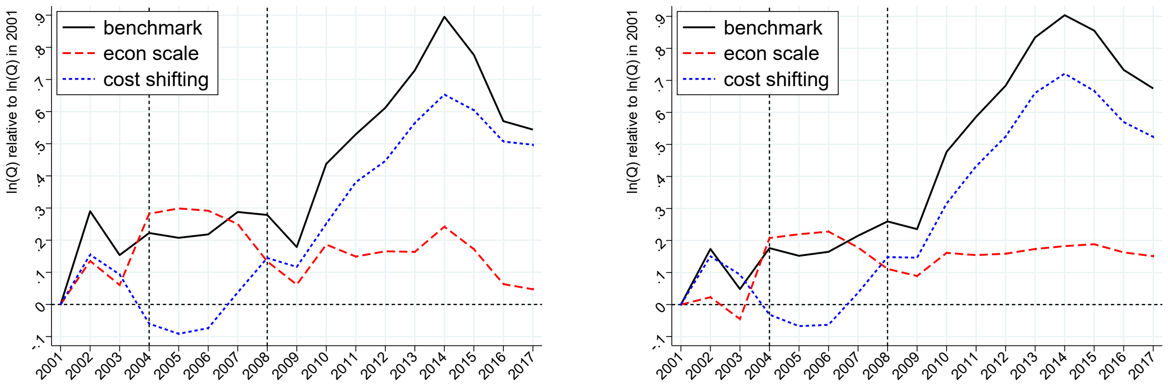
Note: Moving from A to C depicts a change due to cost shifting, which is attributed to heterogeneous productivity while keeping the cost elasticity ϕ unchanged. Moving from C to B depicts a change due to economies of scale.

Table H7: Decomposition of change in output (thousand tonnes)

	(1) 2001–2014	(2) 2001–2004	(3) 2004–2007	(4) 2008–2011	(5) 2011–2014
Panel A: Licensed 0–10.9m					
Total	36.1	3.5	2.8	8.4	22.9
econ scale	8.0	4.3	-0.5	1.0	6.8
cost shifting	28.0	-0.8	3.3	7.3	16.1
Panel B: Licensed 11–14.9m					
Total	89.2	7.3	4.9	26.1	49.1
econ scale	15.3	8.4	0.2	3.6	7.9
cost shifting	73.9	-1.1	4.7	22.5	41.2
Panel C: Licensed 15–20.9m					
Total	280.0	8.6	17.6	73.4	158.9
econ scale	91.1	14.0	8.1	16.7	55.0
cost shifting	188.9	-5.4	9.5	56.7	103.9
Panel D: Licensed 21–27.9m					
Total	403.5	13.4	37.1	94.6	238.4
econ scale	221.8	12.1	26.2	52.7	143.9
cost shifting	181.7	1.3	10.9	41.8	94.6

Table H8: Decomposition of change in $\ln(Q)$

	(1) 2001–2014	(2) 2001–2004	(3) 2004–2007	(4) 2008–2011	(5) 2011–2014
Panel A: Licensed 0–10.9m					
Total	0.90	0.22	0.07	0.25	0.37
econ scale	0.24	0.28	-0.03	0.01	0.10
cost shifting	0.66	-0.05	0.09	0.24	0.27
Panel B: Licensed 11–14.9m					
Total	0.90	0.18	0.04	0.33	0.32
econ scale	0.21	0.22	-0.03	0.05	0.04
cost shifting	0.69	-0.05	0.07	0.27	0.28
Panel C: Licensed 15–20.9m					
Total	1.26	0.11	0.14	0.43	0.42
econ scale	0.51	0.18	0.05	0.13	0.17
cost shifting	0.75	-0.07	0.10	0.29	0.24
Panel D: Licensed 21–27.9m					
Total	1.21	0.15	0.23	0.38	0.40
econ scale	0.64	0.15	0.20	0.20	0.22
cost shifting	0.57	0.00	0.03	0.18	0.17



(a) licensed length 0–10.9m (b) licensed length 11–14.9m

(c) licensed length 15–20.9m (d) licensed length 21–27.9m

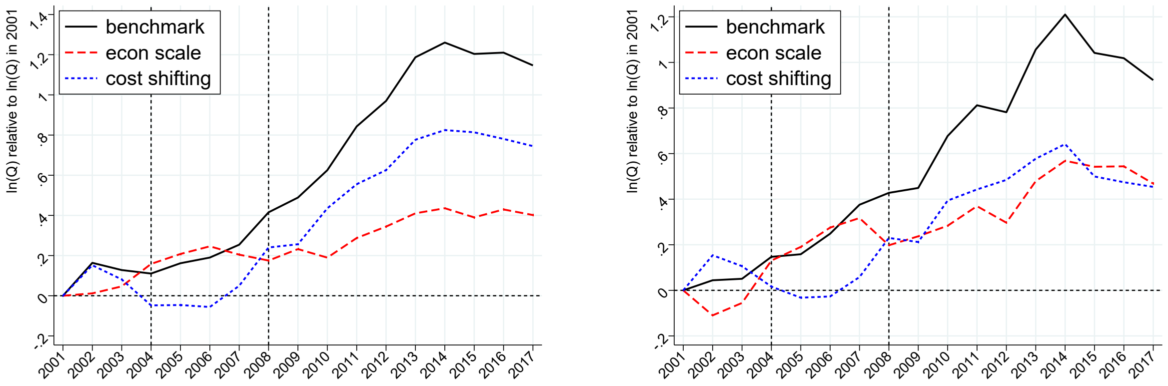
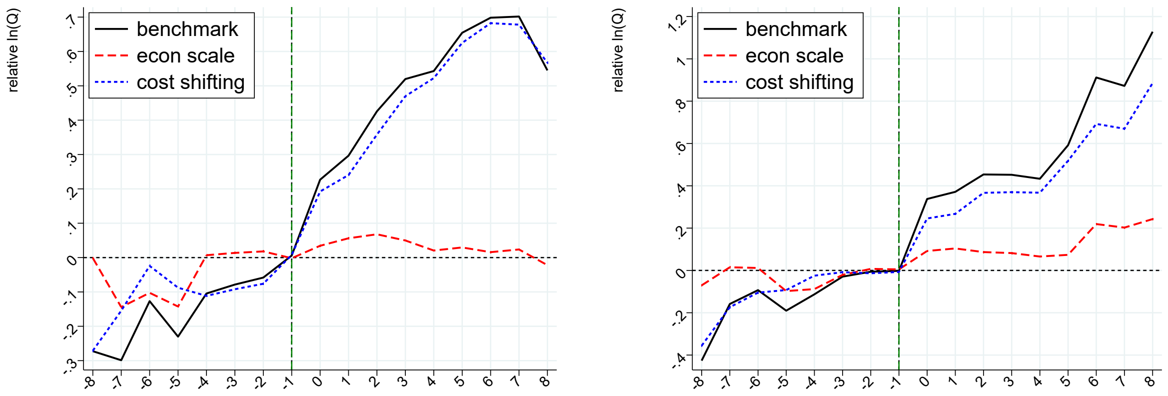


Figure H16: Decomposition of change in $\ln Q_{it}$ (logged thousand tonnes) by year

Note: The figure plots the change in output within a vessel from 2001 to each year. The average change number for each year is weighted by the vessel share of group catch.



(a) licensed length 11–14.9m

(b) licensed length 15–20.9m

(c) licensed length 21–27.9m

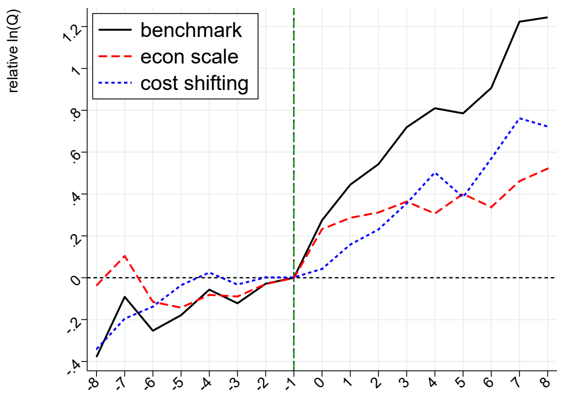
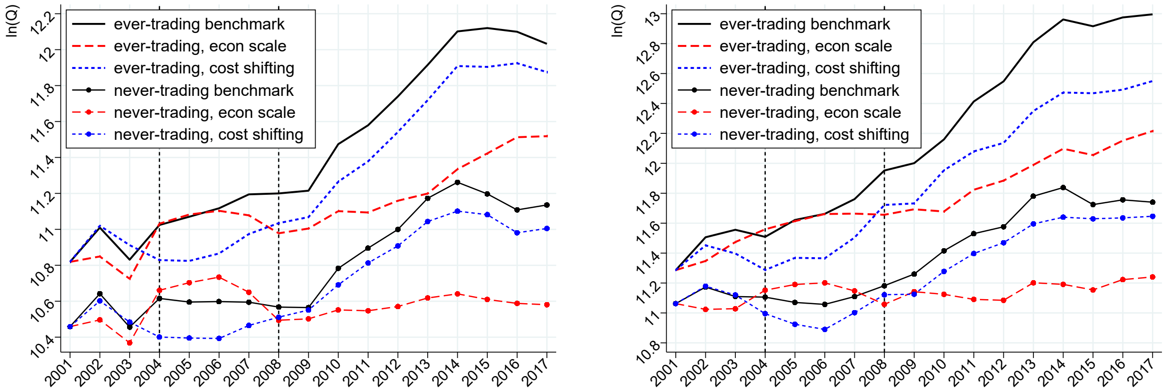


Figure H17: Decomposition of change in $\ln Q_{it}$ (thousand tonnes) by years from the first time a vessel acquires traded quotas

Note: The figure plots the change in output within a vessel over years. All changes are relative to the year when a vessel acquires traded quotas in its first time.


(a) licensed length 11–14.9m

(b) licensed length 15–20.9m

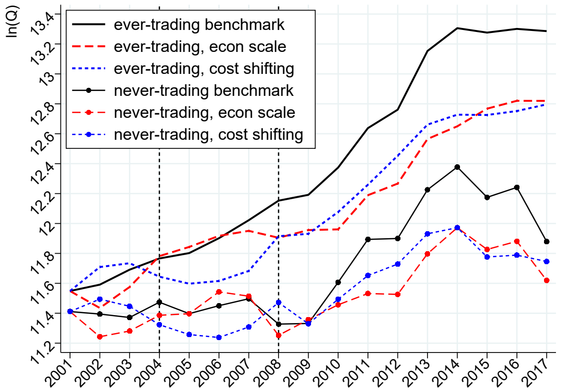

(c) licensed length 21–27.9m

Figure H18: Counterfactual catch by trading action and licensed length group

Note: The figure plots the counterfactual output (thousand tonnes) by quota-acquisition status. The curves with (without) dot markers represents output growths of vessels that never (ever) acquired quotas. The benchmark represents the average output (Q_{it}) weighted by a vessel's share of catch in the licensed length group. The long dash red line shows the evolution of the output as if there were only "economies of scale" effect and the other "cost shifting" effect were 0. The short dash blue line shows the path of the output if the change had been only due to cost shifting.