The value of information in a congested fishery*

Gabriel Englander[†] Larry Karp[‡] Leo Simon[§]

July 21, 2025

Abstract

We model a fishery with potential congestion, in which firms obtain public and private signals about the location of the densest fish stock. We analytically determine the regions of parameter space where greater precision of public and/or private information increases welfare, and we examine the effects of two types of information sharing. Using high-resolution data from the world's largest fishery, we estimate the structural model. Point estimates imply that more precise private information raises welfare, whereas more precise public information has a negligible effect on welfare. Moreover, welfare is much more sensitive to changes in the precision of private information than to changes in the precision of public information. This difference reflects the fact that public information increases congestion more than private information does. We also find empirically that a small amount of information sharing can reduce welfare, whereas more extensive information sharing raises welfare, as do information clubs.

KEYWORDS: value of information, fishery congestion, Peruvian anchoveta

JEL CLASSIFICATIONS: D83, Q22, Q28, Q56, O13

*We thank Philippe Marcoul, Martin Smith, Kate Pennington, and seminar participants at UC Santa Barbara, UC Berkeley, the London School of Economics, University of Florida, University of Hawaii, and Pontificia Universidad Católica de Chile for helpful comments. We greatly benefited from comments by five anonymous referees and the editor. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

†Development Research Group, World Bank. 1776 G St NW, Washington, DC 20006. aenglander@worldbank.org.

‡Department of Agricultural and Resource Economics; University of California, Berkeley. 207 Giannini Hall Berkeley, CA, 94720-3310, USA. karp@berkeley.edu.

§Department of Agricultural and Resource Economics; University of California, Berkeley. We deeply regret that Leo died before this paper's completion. Leo, we will miss you.

1 Introduction

Better information can increase the efficiency of natural resource use, lowering extraction costs by changing the location or timing of production. However, more precise information can cause producers to converge on the same location or increase production at the same time, thereby increasing congestion (Brown, 1974). The increased congestion externality can outweigh the direct benefit of better information. Adapting a model by Angeletos and Pavan (2007) (hereafter, AP07), we consider the value to an industry of three types of changes in information. The first type involves an exogenous increase in the precision of public and/or private information, e.g., arising from better measurement. The second two types involve, respectively, exclusionary and non-exclusionary information sharing. With (exclusionary) "information clubs", groups of agents share their private information among club members, but not with agents in different clubs. Under (non-exclusionary) "global" information sharing, a planner or industry association collects a fraction of agents' private information and amalgamates it to create a more precise public signal, thus making agents' diminished private information less precise. These three changes have different effects on behavior and payoffs.

Our primary contribution is to estimate the parameters of this model in order to assess the effect of the three types of information change in an industry of global importance. A large theoretical literature studies the effect of more precise information, and of information sharing, in settings where their welfare effect may be ambiguous. However, to the best of our knowledge this theory has not previously been confronted by data. The estimation uses high-resolution data from Peru's anchoveta fishery, the world's largest fishery, accounting for 8% of global marine fish catch (FAO, 2018). We find that more precise private information would increase welfare. By contrast, more precise public information has a negligible effect on welfare. This difference reflects the fact that public information increases congestion to a much greater extent, compared to private information. We also find that both information clubs and at least a modest amount of global information sharing raise welfare.

We follow the literature that treats the distinction between public and private information as exogenous; both types of signals are correlated with the payoff-relevant state of nature.[1] All agents receive the public signal, and each receives a private signal. Greater precision of either signal enables agents to tailor their actions to the true state of nature, tending to increase their profits. However, information-induced changes in actions can reduce profits, e.g., by

[1]Myatt and Wallace (2012) propose an alternative formulation where agents obtain signals from n sources and decide how much costly "receiver effort" to devote to interpreting each. As some parameters and the agent's equililbrium effort changes, the signals vary from private (uncorrelated across agents) to public (perfectly correlated across agents). An exogenous distinction between public and private signals is appropriate in our setting because firms have (intrinsically) private information from their own past activities and public information from satellites.

increasing congestion. Bergemann and Morris (2013, p. 1253) note that a more precise public signal leads to a substantial increase in the correlation of actions across agents, but a small increase in the correlation between an agent's action and the state of the world. In our setting, this means that a more precise public signal leads to a large increase in congestion, while enabling fishers to get only slightly closer to the most favorable fishing ground. By contrast, a more precise private signal leads to higher correlation between the agent's action and the state of the world (getting closer to the ideal fishing ground), but greater dispersion of agents' actions (less congestion). Therefore, increased precision of private information is more likely to improve efficiency. In our setting, global information sharing reallocates information in the "wrong" direction, making the public signal more precise and the private signal less precise. Exclusionary information clubs create a new type of signal and also change agents' decision problem: with the same information as other club members, an agent can predict their actions.

AP07 provides our starting point, but our focus and analytic results differ. AP07 compare the noncooperative equilibrium with the outcome under a team. Agents in the noncooperative equilibrium choose actions to maximize their own welfare. In the team setting, they choose actions to maximize aggregate welfare, but without being able to share information. Because our research question does not involve the team problem, we can extend the domain of parameter space to include greater congestion costs. We use a special case of their model, relevant to the fishing context, where the noncooperative equilibrium is inefficient only under incomplete information. By focusing on this case, we are able to provide more definitive comparative static results.

An earlier literature, motivated largely by antitrust considerations, studies information sharing. The collection of results in Clarke (1983), Vives (1984) and Gal-or (1986) show that whether firms benefit from sharing their private information depends on whether competition is Cournot or Bertrand and on whether the unknown parameter affects demand or costs. In an $n-$agent fishing model with congestion and uncertain stock size, Marcoul (2020) identifies the region of parameter space where industry profits are higher when all firms make their private information public; his numerical examples illustrate the case with two asymmetric exclusionary information clubs. Vives (1990) uses a continuum-of-agents monopolistic competition model in which a voluntary trade association collects private information from participating firms. All firms choose to participate if only participants receive the collected information; no firms participate in the non-exclusionary setting where all firms receive the collected information regardless of their participation. Calvó-Armengol, De Martí, and Prat (2015) also endogenize information sharing by studying how the formation of communication networks between agents influences equilibrium actions. We extend this literature by allowing the industry to split into an arbitrary number of symmetric clubs with exclusionary information sharing; and, under global (non-exclusionary) information sharing, allowing the trade association/planner to collect only a fraction of each firm's private information.

The strands of the literature that study the effect of more precise information or information sharing use similar models; both changes alter the moments of agents' signals. Whereas improved measurement can in principle raise the precision of one signal without altering the other moment(s), under information sharing the moments of the two signals are inextricably linked. Under global information sharing, the transfer of some private information to the public increases the precision of the new public signal and lowers the precision of the now-degraded private signal. Exclusionary information clubs replace the private signal with a club signal, changing agents' decision problem by giving them the same information as other club members. Nevertheless, the similarity of the equilibrium structure with and without information sharing leads to a unified analysis.[2]

Morris and Shin (2002) study a version of the Keynesian beauty contest in which an agent's payoff increases as their decision is more closely aligned to the true market fundamental, and also closer to other agents' decisions: the opposite of congestion costs. Ui and Yoshizawa (2015) classify linear-quadratic-Gaussian games into eight types, distinguished by how welfare changes with increased precision of public and private information. Our model corresponds to their Type +III game, where an increase in the precision of public information improves welfare only if greater precision of private information also does. Our analysis also relates to a literature on strategic experimentation, in which congestion (coordination) is a negative externality because it reduces the opportunity to learn from others (Bolton & Harris, 1999; Callander & Harstad, 2015). Unlike in those papers, the public signal in our context is exogenous.

Section 2 explains how features of the anchoveta fishery map into our model assumptions. Section 3 lays out the model and provides formulae for equilibrium actions and payoffs. The following section describes the comparative statics of welfare with respect to exogenous changes in signals' precision; we then discuss information clubs and global information sharing. Section 5 describes the data, and the next two sections estimate the model parameters. Section 8 uses our point estimates and analytic results to evaluate the three types of changes in information. Section 9 discusses extensions and a final section concludes.

[2] If more precise signals or the reallocation of information (arising from either global information sharing or information clubs) lower welfare, no effort would be spent in these endeavors. When any of these changes raise welfare, we would need to know the cost of achieving the change to predict the equilibrium level of the change. Myatt and Wallace (2012) model the cost of increasing the precision of received signals, and Calvó-Armengol, De Martí, and Prat (2015) model the cost of increasing the precision of both received and sent signals. We do not attempt to extend our model in that direction.

2 Institutional context

Here we explain how the institutional features of the Peruvian anchoveta fishery motivate the characteristics of our model.

Congestion without inter-seasonal stock externalities. Congestion may be important in the fishery because more intense nearby fishing depletes the local anchoveta population and reduces vessels' ability to maneuver and deploy nets. These congestion externalities may increase fishing costs. In contrast, inter-seasonal stock externalities are relatively unimportant in this fishery because of the regulator's ability to (i) limit catch, (ii) respond to new information to protect the stock, and (iii) promote economically efficient fishing.[3]

(i) The regulator (PRODUCE) consults with the marine science agency (IMARPE) to set an industry-wide limit on the total tons that can be landed each season, called the total allowable catch (TAC). Tons "landed" refers to vessels' transfer of their catch to a processing plant. The regulator sets the TAC before the start of the season. There are two fishing seasons per year for Peru's North-Central anchoveta stock, each lasting about three months.[4]

(ii) In four out of the six seasons in our data, the TAC announced at the beginning of the season is binding. In the second season of both 2017 and 2019 the regulator closed the fishing season early, before the TAC was reached, due to the detection of significant spawning activity or the presence of juvenile anchoveta (Englander, 2023).

(iii) Individual vessel quotas (IVQs) divide the TAC among vessels. The IVQ regulation precludes entry of new vessels into the fishery and gives firms the incentive to minimize costs, thereby avoiding overcapitalization. Each vessel is entitled to land a fixed percentage of the TAC each season.

The location of catch does not affect the TAC or IVQs. There are 17 main ports in the North-Central fishery at which vessels can land their catch. Landings data for 2016 show that 47% of vessels' landings occur at their most common port (PRODUCE, 2020c). IVQs may be transferred among vessels owned by the same firm. But to transfer an IVQ across firms, the vessel itself must be sold (Natividad, 2016).

Because entry is restricted and aggregate catch is fixed by regulation, the antitrust considerations that motivated the literature on information sharing are not relevant in our setting (Vives, 1990). To the extent that regulatory limits on catch are set efficiently, they solve the stock externality. Without an offsetting distortion, congestion always lowers efficiency because it increases the cost of harvesting a given level of catch.

[3] Stock externalities would likely decrease the value of public information.

[4] Peru's North-Central stock occurs entirely within Peruvian jurisdiction. The stock ranges from Peru's northern boundary to the 16th parallel south. Peru's Southern anchoveta stock is shared with Chile and is subject to different regulations. The North-Central stock accounts for 95% of tons caught per year (Englander, 2023). This paper analyzes data from the North-Central fishery only.

Huang and Smith (2014) and Sanz and Diop (2021) provide empirical evidence of the importance of congestion in shrimp fisheries. Huang and Smith note that greater congestion can improve efficiency by partially offsetting a stock externality: excessive harvest that reduces the future stock. Unlike Huang and Smith (2014), we estimate only spatial congestion, not a dynamic stock-related externality.

Increasing information precision and policy implications. Existing and emerging technologies make the findings of our analysis actionable. For example, firms can increase the precision of their private signals by deploying acoustic buoys to estimate anchoveta biomass in specific locations (Brehmer et al., 2019; Simmonds et al., 2009). New satellites, such as the Plankton Aerosol Cloud ocean Ecosystem (PACE), can increase the precision of public information by providing better measures of the correlates of anchoveta abundance (National Aeronautics and Space Administration, 2024). More precise information could help firms' vessels reach their quota at lower cost, for example by reducing search costs. Search comprises 20% of time spent during fishing trips (Joo et al., 2015).

Payoff function. Our model omits fish price because it is exogenous. With a binding constraint on landings, the firm's seasonal revenue is exogenous, so variations in firm behavior affect only costs. Accordingly, our payoff function focuses on cost rather than revenue.

All vessels receive the same price for their catch; the price per ton is a fixed fraction of the price of fishmeal in Hamburg (Fréon et al., 2014; Hansman et al., 2020). Thus, more precise information will not change prices, unlike what occurred in South Indian fisheries (Jensen, 2007). According to a collective bargaining agreement with the seven largest firms, vessels that land anchoveta that will be processed into fishmeal and fish oil receive 1.792% of the Free On Board price of fishmeal per ton of anchoveta landed (SUPNEP, 2017).[5] Under the collective bargaining agreement, a vessel's crew shares revenue in fixed proportions: the captain receives "two parts" (twice as much as a regular fisher), the second-in-command and first engineer receive one and a half parts, and regular fishers receive one part (Englander, 2023; SUPNEP, 2017).

Information sharing within firms. Firms promote information sharing among their own vessels, while fishers tend to avoid revealing the locations of productive fishing grounds to others (Marcoul, 2020; Welch et al., 2022). For example, vessels belonging to the same firm transmit fishing locations and tons caught to firm headquarters. The two firms visited by one of the authors in December 2019 had central "command centers" in which staff receive such data from individual vessels, aggregate the information, and send daily reports with suggested fishing locations to their vessels.

Global information sharing. Firms are required to send their private information

[5] 97% of tons landed of anchoveta are processed into fishmeal and fish oil (PRODUCE, 2018).

about the size and location of their catch to the regulator. Two compulsory monitoring technologies allow the regulator to verify the information submitted (Section 5.3). However, the regulator does not currently make this information publicly available, so there is scope for increased global information sharing. Additional private information could be costly to collect or difficult to verify. Therefore, our model's feature that allows the regulator to collect and amalgamate a fraction (rather than the entirety) of firms' private information to create a more precise public signal is important for model realism.

Information clubs. Firms are reluctant to share information with others, making information clubs difficult to form, and likely limiting their size. However, Gatewood (1984) provides examples where fishers sometimes share information about stock location. Therefore, rather than a single club consisting of all firms, our model of small information clubs, in which firms reveal all of their private information to fellow club members, is also important for model realism.

3 The model

Each of $n < \infty$ firms receives a public and a private signal about θ, the location of the "ideal fishing ground", where the stock is densest. To streamline the exposition, we begin with the situation where $\frac{n}{c}$ equal-sized information clubs form, each having c members. Members of an information club share their private information, but act noncooperatively in choosing their location. Lemma 1 reports the Bayesian Nash equilibrium decision rule and payoff in the game with $c \geq 1$. Specializing to $c = 1$, we obtain the case of primary interest, where firms do not share information. Apart from their receipt of a private signal, firms are homogeneous; the empirical implementation relaxes this assumption.

Firms face either congestion costs or benefits from operating close to each other. If there are congestion costs, each firm wants to be far from others; with positive externalities, the firm wants to be close to others. Using the totality of their information, the firm estimates the ideal location and forms beliefs about other agents' locations. The firm then decides where on the real line to locate, balancing the benefit of being close to the ideal fishing ground with the costs or benefits of being close to others.

Firms have a diffuse prior on θ and receive the public signal $y = \theta + \varepsilon_y$, with $\varepsilon_y \sim N\left(0, \sigma_y^2\right)$, and a private signal $x_i = \theta + \varepsilon_{x_i}$, with $\varepsilon_{x_i} \sim N\left(0, \sigma_x^2\right)$. The noise across agents' private signals is uncorrelated: $\mathbb{E}\varepsilon_{x_i}\varepsilon_{x_j} = 0$, for $i \neq j$. However, the noise in an agent's private signal may be correlated with the noise in the public signal: $\mathbb{E}\varepsilon_y\varepsilon_{x_i} = \rho\sigma_y\sigma_x$.[6] The covariance matrix

[6]Public signals use satellites covering the entire fishery, producing daily environmental data across fishing locations (Section 5). In contrast, private signals are specific to where a firm's vessels fished the previous day.

for the public signal and the n private signals is positive definite if and only if $n\rho^2 < 1$ (Online Appendix B). The number of firms actively fishing in the anchoveta industry varies throughout the season; the median number active per day in our sample is $n = 135$. With this value, positive definiteness requires $|\rho| < 0.09$. Our results are substantially the same for $\rho = 0$ and $\rho = 0.09$, so in the interest of simplicity, subsequent sections set $\rho = 0$. This section retains ρ as a parameter because future empirical work involving smaller industries may find the general formulae useful.[7]

Define the ratio of standard deviations as $r \equiv \frac{\sigma_y}{\sigma_x}$; a larger r corresponds to a relatively less precise public signal. Based on the public and only their own private signal, firm i's Bayesian posterior for θ is $\delta y + (1-\delta)x_i$, with

$$\delta \equiv \frac{1-\rho r}{1+r^2-2\rho r}. \tag{1}$$

This weight is also the Best (minimum variance) Linear Unbiased Estimator (BLUE). We refer to δ as the BLUE weight to distinguish it from the Bayesian Nash equilibrium weight.[8]

We now consider the non-public information shared by members of a club. Because all private signals are equally informative and their noise is uncorrelated, the average of these signals is a sufficient statistic for the set of the club's signals. The average of non-public signals for club s is $\tilde{x}_s \equiv \frac{1}{c}\sum_{i\in s} x_i$. Similarly, we define $\tilde{\varepsilon}_s = \frac{1}{c}\sum_{i\in s}\varepsilon_{x_i}$, the average error of club members' uncorrelated private signal noise. Given our assumptions on x_i[9]

$$\tilde{x}_s \sim N\left(\theta, \frac{\sigma_x^2}{c}\right) \text{ and } cov\left(\tilde{x}_s, y\right) = \mathbb{E}_{\varepsilon_y,\{\varepsilon_{x_i}\}}\left(\varepsilon_y \frac{\sum \varepsilon_{x_i}}{c}\right) = \rho\sigma_x\sigma_y.$$

The correlation of $\tilde{x}_s$ and y is therefore $\rho\sqrt{c}$. The club's collective private signal, $\tilde{x}_s$, is more precise and (weakly) more highly correlated with the public signal, compared to any club member's individual private signal, x_i.

Noise is potentially correlated because private signals are derived from a subset of locations from the public signal data—they share common information from overlapping locations. However, the noise in different firms' private signals is uncorrelated because firms' vessels physically cannot fish in the same places and times.

[7]Our unconstrained point estimate of ρ exceeds 0.09. A previous working paper showed that results are similar with the larger point estimate and with $\rho = 0$. Footnotes and Online Appendices discuss $\rho \neq 0$.

[8]δ is positive unless $\rho r > 1$. When the noise in public and private signals is highly positively correlated and in addition the public signal is relatively noisy (so that ρr is large), the public signal contains a great deal of noise relative to its additional information. In that case, the BLUE gives negative weight to the public signal. The weight on the private signal is positive ($1-\delta > 0$) unless $\rho > r$. When the noise in public and private signals is highly correlated and r is small, the private signal contributes a great deal of noise relative to its additional information, so the BLUE gives that signal negative weight.

[9]The notation $\mathbb{E}_{\varepsilon_y,\{\varepsilon_{x_i}\}}$ means that the expectation is taken over ε_y and over the c random variables $\{\varepsilon_{x_i}\}$. Where the meaning is clear, we suppress these subscripts.

If nature chooses θ, firm i chooses location k_i, and $j \neq i$ chooses location k_j, i's payoff is

$$J(k_i|\{k_j\}, \theta) \equiv \frac{1}{n-1}\frac{B}{2}\sum_{j\neq i}(k_i - k_j)^2 - \frac{A}{2}(k_i - \theta)^2. \tag{2}$$

The parameters A and B are constant, with $A > 0$ and $B \gtreqless 0$.

The second term on the right side of Equation 2 captures the increased costs arising from greater distance between an agent and θ. Fishing farther from where the stock is densest may reduce revenue for a particular trip, or require that the trip be extended, increasing costs; it might also increase the number of future trips needed to reach the seasonal quota. If fish are spatially diffuse, a firm's position relative to the ideal location matters little, and $A \approx 0$. However, if fish are concentrated near the ideal location, the penalty of being distant from it is significant; here, A is large. The magnitude of A also depends on the total biomass relative to fishers' capacity. If the biomass is very large and the capacity of a firm to catch fish is limited, the distance from the ideal location might not be important, making A small.

The first term in $J(k_i|\{k_j\}, \theta)$ captures the costs or benefits associated with *dispersion*, the distance between agent i and the average location of other agents. For $B > 0$ there is congestion: agent i benefits from being far from other agents due to lower fishing costs. Here, its payoff increases with the average distance between it and other agents. For $B < 0$, proximity to other agents benefits i, e.g., due to increased safety.[10] We define the normalized measure of congestion $\tau \equiv \frac{B}{A}$.

With linear strategies, i's location is $k_i = \gamma y + (1-\gamma)\, x_i$, and i believes that agent $j \neq i$ chooses $k_j = \eta y + (1-\eta)\, x_j$. The symmetric Bayesian Nash equilibrium weight, $\gamma^{NE,c}$, when each club has c members, satisfies $\gamma = \eta$. Appendix A contains sketches of all proofs and Online Appendix B.1 contains the detailed proofs. The model parameters satisfy:[11]

Assumption 1. *(i) $A > 0$, (ii) $\tau < 1$, (iii) $n\rho^2 < 1$.*

Lemma 1. *Under Assumption 1, when the n agents coalesce into T information clubs, each with $c \geq 1$ members, the Bayesian Nash equilibrium weight on the public signal is*

$$\gamma^{NE,c} = \frac{1 - \frac{n-c}{n-1}\tau - cr\rho}{1 - \frac{n-c}{n-1}\tau + cr^2 - 2cr\rho}, \tag{3}$$

[10]The quadratic form is an approximation of a more general function that increases with proximity to θ and increases or decreases (depending on the sign of B) with dispersion. The quadratic function is consistent with the assumption that biomass density decreases monotonically with the distance from θ, and catch increases monotonically with the density. We confirm that this relationship between expected payoff and distance to the ideal location holds in our data.

[11]We use a larger parameter space than AP07 ($\tau < 1$ instead of $\tau < 0.5$) but we cannot use their adaptation of Morris and Shin's (2002) uniqueness proof over $0.5 \leq \tau < 1$; there we simply assume that the equilibirum is linear in information. Appendix A discusses these points.

and the equilibrium payoff is

$$P = \tfrac{A}{2}\sigma_x^2 S\left(\tau, \gamma\left(r, \tau, \rho, c\right), r, \rho, c\right) \text{ with } S\left(\tau, \gamma\left(r, \tau, \rho, c\right), r, \rho, c\right) \equiv$$
$$2\tfrac{1}{c}\left(\tfrac{n-c}{n-1}\right)\left(1-\gamma\right)^2 \tau - \left[\gamma^2 r^2 + \left(1-\gamma\right)^2 \tfrac{1}{c} + 2\gamma\left(1-\gamma\right)\rho r\right]. \tag{4}$$

evaluated at $\gamma = \gamma^{NE,c}$*, the equilibrium decision rule.*

4 Analysis

Propositions 1 and 2 in Section 4.1 examine the welfare effect of an exogenous change in precision, e.g., arising from better measurement. The next two subsections study changes in signals' moments arising from the transfer of information. Section 4.2 uses Lemma 1 to determine firms' incentives to form information clubs. Section 4.3 studies "global information sharing", where a regulator extracts a portion of each agent's private information and amalgamates it into a more precise public signal, thereby changing all of the moments. Using formulae for this moment transformation, we can apply Lemma 1 to determine the equilibrium effect of global information sharing.

By inspection of Equations 3 and 4, the decision rule and the payoff per firm are invariant to n if there are no information clubs ($c = 1$). This invariance is due to the fact that an agent's payoff depends on the average deviation between its location and all other agents' locations. That average does not change with the number of agents, for a given decision rule. Sections 4.2 and 4.3 explain why the equilibrium depends on n with clubs or global information sharing.

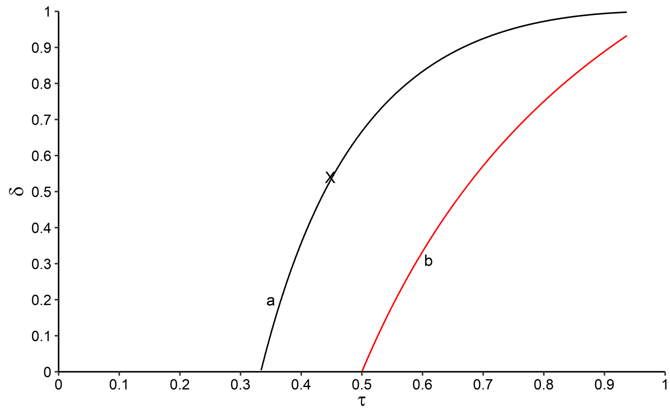
Under global information sharing the regulator might collect only a fraction of firms' private information (to make the public signal more precise), but we assume that club members transfer all their private information to fellow club members. This assumption helps to streamline the exposition by avoiding repetition. It also means that every club member has two types of signal: club and public. If we had instead assumed that firms transfer only a fraction of their private information to fellow club members, firms would have three types of signal: private, club, and public. That more general setting would not permit a simple nesting of the cases with and without clubs.

4.1 Exogenous change in precision

Setting $c = 1$ and $\rho = 0$ in Equation 3, we obtain the Bayesian Nash equilibrium weight

$$\gamma_{|\rho=0}^{NE,1} = \delta\frac{1-\tau}{1-\tau\delta} < \delta \text{ iff } \tau > 0. \tag{5}$$

Figure 1: Boundaries that determine the welfare effect of more precise information



Notes: For $\rho = 0,\ c = 1$: (i) more precise public information increases welfare if and only if (τ, δ) lies above the curve labeled a; (ii) more precise private information increases welfare if and only if (τ, δ) lies above the curve labeled b. The point X identifies our preferred point estimate $(\tau, \delta) = (0.448, 0.54)$, where an increase in the precision of either signal raises welfare.

When $\tau = 0$, a firm's payoff is unaffected by other agents' locations. In this limiting case, $\gamma_{\tau=0}^{NE,1} = \delta$: firms use the BLUE weight on the public signal. However, with congestion ($\tau > 0$) the equilibrium weight on the public signal is smaller than, and an increasing function of, δ.

The next two propositions discuss the welfare effect of increasing the precision of one signal, holding the precision of the other signal fixed. Figure 1 provides an aid in reading these propositions. Proposition 1 uses the function $\delta^a(\tau) \equiv \frac{3\tau-1}{\tau^2+\tau}$, shown as the curve labeled a in Figure 1. Proposition 2 uses the function $\delta^b(\tau) \equiv \frac{2\tau-1}{\tau}$, shown as the curve labeled b. An increase in the precision of the public signal corresponds to a reduction in r and an increase in δ: a movement to the north in the figure. An increase in the precision of the private signal corresponds to an increase in r and a decrease in δ: a movement to the south in the figure.

Proposition 1 considers the welfare effect of a change in the precision of the public signal, holding the precision of the private signal fixed.

Proposition 1. *Under Assumption 1 with $\rho = 0$ and $c = 1$: (i) For $\tau < \frac{1}{3}$, welfare increases with the precision of the public signal. (ii) For $\tau > \frac{1}{3}$ welfare is minimized with respect to the precision of public information on the curve $\delta^a = \frac{3\tau-1}{\tau^2+\tau}$. For δ above this curve, welfare increases with a more precise public signal, and below the curve welfare increases with a less precise public signal. (iii) If it were possible to change the precision of the public signal costlessly, the optimal choice is bang-bang: for $\tau < 0.5$ it is optimal to make the public signal infinitely precise. For $\tau > 0.5$, where congestion is severe, it is optimal to eliminate the public signal.*

Proposition 2 considers the welfare effect of a change in the precision of the private signal, holding the precision of the public signal fixed.

Proposition 2. *With Assumption 1 and $\rho = 0$ and $c = 1$: (i) For $\tau < 0.5$, welfare increases with the precision of the private signal. (ii) For $\tau > 0.5$ welfare is maximized with respect to the precision of the private signal on the curve $\delta^b = \frac{2\tau-1}{\tau}$. For δ above this curve, increased precision of the private signal raises welfare. Below this curve, decreased precision of the private signal increases welfare.*

Because the curves labeled a and b in Figure 1 are monotonically increasing, the higher is the relative precision of the public signal (the larger is δ), the larger is the critical level of congestion above which more precise (public or private) information lowers welfare. In addition, because the curve b lies to the right of the curve a, a smaller amount of congestion is required to make increased precision of the public (compared to the private) signal harmful.

Greater precision of either the public or private signal increases the accuracy of agents' information, i.e., it reduces the variance of i's forecast error, $v_i \equiv \theta - \mathbb{E}[\theta|y, x_i]$. AP07 note that the correlation between agents' forecast errors is $corr(v_i, v_j) = \delta$. When neither signal is perfectly precise, an increase in the precision of the public signal increases this correlation, whereas an increase in the precision of private information reduces it.[12] That is, a more precise private signal increases informational heterogeneity, causing firms to locate farther apart; and a more precise public signal lowers this heterogeneity, causing firms to move closer together. This difference explains why for the region of parameter space between the curves labeled a and b in Figure 1, increased precision of the private signal raises welfare, but increased precision of the public signal lowers welfare.[13]

A more precise public signal tends to benefit agents by enabling them to get closer to the target. However, with agents moving closer to the target they also tend to move closer to each other. For $\tau > 0$, the reduction in dispersion harms agents, so the net effect of a more precise public signal is ambiguous. For large τ (to the right of curve a) congestion is important enough that increased precision of the public signal lowers welfare. The increased precision of the private signal creates similar opposing forces. However, a more precise private signal creates a smaller reduction in dispersion compared to the more precise public signal.[14]

[12] If either signal is perfectly precise, all firms know θ and there is no informational heterogeneity.

[13] AP07 obtain either necessary or sufficient conditions (but not both) to sign the welfare effect of more precise public or private information for $0 < \tau < 0.5$. Our ranking criteria is both necessary and sufficient and it holds for $\tau < 1$. In addition, we obtain the welfare effect of non-marginal changes in the precision of both signals. We are able to obtain these stronger results because we use an explicit expression for the payoff in terms of model primitives. For $\tau < 0$, where agents benefit by being close to other agents, we reproduce AP07's results.

[14] Online Appendix B.2.2 shows how the boundaries in Figure 1 change when $\rho > 0$. Even a large value of ρ causes almost no change to the boundary b. That is, even a substantial correlation between noise in the public and private signals has almost no effect on the comparative statics of welfare with respect to the precision of private information. For τ in the range that is plausible for the anchoveta fishery (where likely

There is a simple relation between the elasticities of welfare with respect to public and private information. Parameter changes can flip the sign of the payoff (P, given by Equation 4), but this sign is of no intrinsic interest. We therefore use $|P|$ instead of P in the denominator of the elasticities, so that these have the same sign as the derivatives. With this convention, the sum of the two elasticities equals 2. The elasticity with respect to the precision of public information is

$$\begin{aligned}\alpha \equiv \frac{dP}{d\left(\frac{1}{\sigma_y}\right)}\frac{\left(\frac{1}{\sigma_y}\right)}{|P|} = -\frac{dP}{d(\sigma_y)}\frac{\sigma_y}{|P|} = -\frac{dS}{d\sigma_y}\frac{\sigma_y}{|S|} = -\frac{dS}{dr}\frac{r}{|S|} = \\ sign(-P)2(1-\tau)\frac{2\tau-\tau^2+3r^2\tau-r^2-1}{(-\tau+r^2+1)(2\tau-\tau^2+2r^2\tau-r^2-1)}.\end{aligned} \tag{6}$$

The equalities in the first line of Equation 6 hold for general ρ, and the second line specializes to $\rho = 0$. Online Appendix B.2.3 provides details. The elasticity of the payoff with respect to precision of private information is

$$\beta \equiv \frac{dP}{d\left(\frac{1}{\sigma_x}\right)}\frac{\frac{1}{\sigma_x}}{|P|} = -\frac{dP}{d\sigma_x}\frac{\sigma_x}{|P|} = sign(-P)\cdot 2 - \alpha. \tag{7}$$

For $P < 0$, as holds for our point estimates, $\beta = 2 - \alpha$. Increased precision of public and private information have the opposite effect on r, which appears in the function S in Equation 4. This reversal accounts for the $-\alpha$ in the definition of β. In addition to this indirect effect operating via r in the function S, increased precision of private information lowers the coefficient $\frac{A}{2}\sigma_x^2$ in the payoff P in Equation 4; when $P < 0$ (so $S < 0$) this lower coefficient increases the payoff.

4.2 The incentive to form clubs: $c \geq 1$

Without knowing the cost of forming information clubs, we cannot determine the equilibrium club size. However, we can show how a change in the club size changes the equilibrium decision rule and changes welfare when there are n firms. In our empirical setting, we find that a larger club decreases the weight on the public signal and raises welfare (Section 8).

Firms know that all members of their club have the same information, so in a symmetric equilibrium they choose the same location. Therefore, fellow club members do not contribute to the dispersion term in the firm's payoff on the right side of Equation 2. As c increases, with $c > 1$, the change in the payoff weight on dispersion relative to the weight on missing

$\tau < 0.65$), the curve a shifts up with $\rho > 0$. That is, positive correlation between the noise in the signals reduces the parameter space over which more precise public information raises welfare.

the target (τ) therefore depends on n, causing the equilibrium to depend on n. For example, if $c = 10$ and $n = 20$, a firm's incentive to increase dispersion (for $\tau > 0$) remains, because there is a chance of being far from members of the other club. In contrast, with $c = 20 = n$, all firms know that equilibrium dispersion is zero, so each firm's only goal is to get close to the ideal location. For given c, the larger is n, the smaller is the change in payoff weights on dispersion due to a club having $c > 1$ members.[15]

4.3 Global information sharing

Here we consider the situation where a regulator (or industry association) extracts from each firm a fraction of their private information. The regulator combines this newly acquired information with the existing public information to create a more precise public signal. Each firm is left with less private information, so their post-transfer private signal is less precise. To study this type of information transfer we modify the model introduced in Section 3 by assuming that each firm receives m private signals, instead of a single private signal as before. The noise in these m private signals is uncorrelated with the noise in other private signals, both for the same firm and across firms. We choose the moments of these private signals so that the aggregate amount of information is unchanged. That is, absent information transfer, the new model is equivalent to our original model.

Ignoring the integer constraint, we assume that the regulator extracts the fraction f of each firm's private information. The new information set for the regulator consists of the original public signal and the information extracted from the n firms. We define the function

$$\varpi \equiv \frac{nr^2}{nr^2 + f^{-1}}, \tag{8}$$

the fractional reduction in the variance of the public signal achieved by the information transfer. We use this function to express the relation between the original and the post-transfer moments: $\{\sigma_y^2, \sigma_x^2\}$ and $\{\sigma_y^{2\prime}, \sigma_x^{2\prime}\}$, respectively. The post-transfer moments are[16]

$$\sigma_y^{2\prime} = (1 - \varpi)\, \sigma_y^2;\ \sigma_x^{2\prime} = \frac{1}{1 - f}\sigma_x^2;\ r' = r\sqrt{(1 - \varpi)(1 - f)} \leq r. \tag{9}$$

The first equality in Equation 9 uses the definition of ϖ; the second follows from the fact that the transfer reduces precision of firms' signals by $f \times 100\%$; the third uses the first two equalities and the definition of r. Greater information sharing (higher f) increases ϖ and σ_x and reduces σ_y and r. Section 8 uses our parameter estimates to assess the welfare effect of

[15] This relation is apparent in Equation 19 in Online Appendix B.1.

[16] Online Appendix B.4 derives Equations 8 and 9 for the general case $\rho \gtreqless 0$.

global information information sharing.

5 Data

The value of more precise public or private information depends on the relative precision of the two signals, r, and on the severity of congestion relative to the importance of being close to where fish are densest, τ. Here we describe the data used to estimate these parameters.

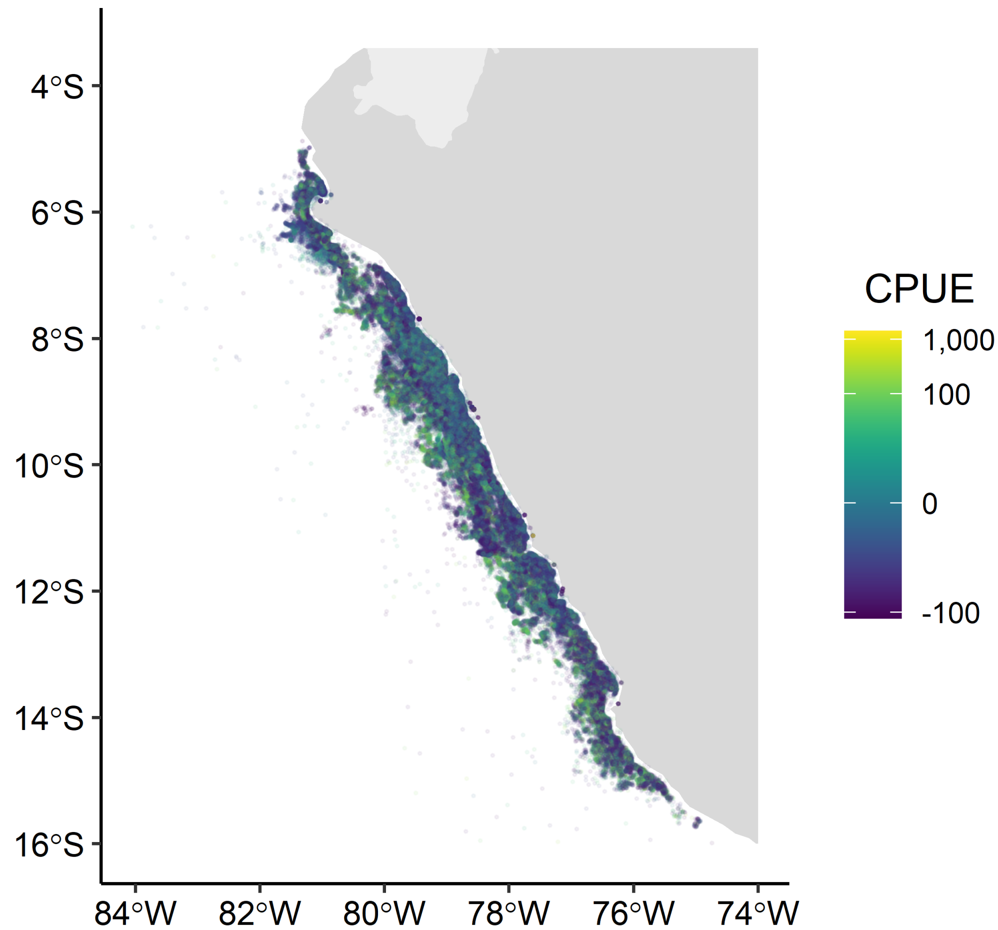
We observe the catch of all 806 industrial vessels in Peru's North-Central anchoveta fishery every time they "set" their net in the water. The unit of observation in this "electronic logbook" data is a set, a vessel-level fishing operation, which occurs at a specific time, longitude, and latitude. Our electronic logbook data spans the six fishing seasons of 2017, 2018, and 2019, containing 246,920 sets (PRODUCE, 2020a).

5.1 Catch per unit effort (CPUE)

We construct a measure of Catch Per Unit Effort (CPUE) using vessel and firm characteristics. The simplest measure of CPUE, tons per set, is one measure used by Peru's marine science agency (IMARPE, 2017). We adjust tons per set by vessel characteristics to account for the fact that sets by larger and more powerful vessels require more energy than sets by smaller and less powerful vessels. We further adjust tons per set by firm characteristics to account for other cost differences across firms. For example, vessels belonging to larger firms may have lower cost per set because they are more technically efficient than vessels that belong to smaller firms. The firm-dependence of our CPUE measure also accounts for the fact that larger firms have more information; that informational advantage is similar to an advantage in equipment. Finally, firm characteristics adjust for the possibility that crews of smaller firms divide revenue differently than crews of the seven largest firms (Section 2).

We construct CPUE by regressing tons per set on the length (in meters) of the vessel, the engine horsepower of the vessel, the gross tonnage of the vessel (a measure of internal volume), the number of vessels owned by the vessel's firm, and indicators for each of the following three firm types. Seven large firms each own at least 19 vessels; 331 "singleton" firms own only one vessel (PRODUCE, 2020b). The remaining vessels belong to 131 medium-sized firms that own between 2 and 10 vessels. The residuals from this regression are our primary measure of CPUE because they condition catch on fishing cost. Figure 2 plots CPUE by location. Repeating our analysis using a simpler measure of CPUE (tons per set minus vessel-level average tons per set) or a measure of CPUE that adjusts for travel costs produces similar parameter estimates (Online Appendix C).

Figure 2: Peruvian electronic logbook data, 2017 to 2019



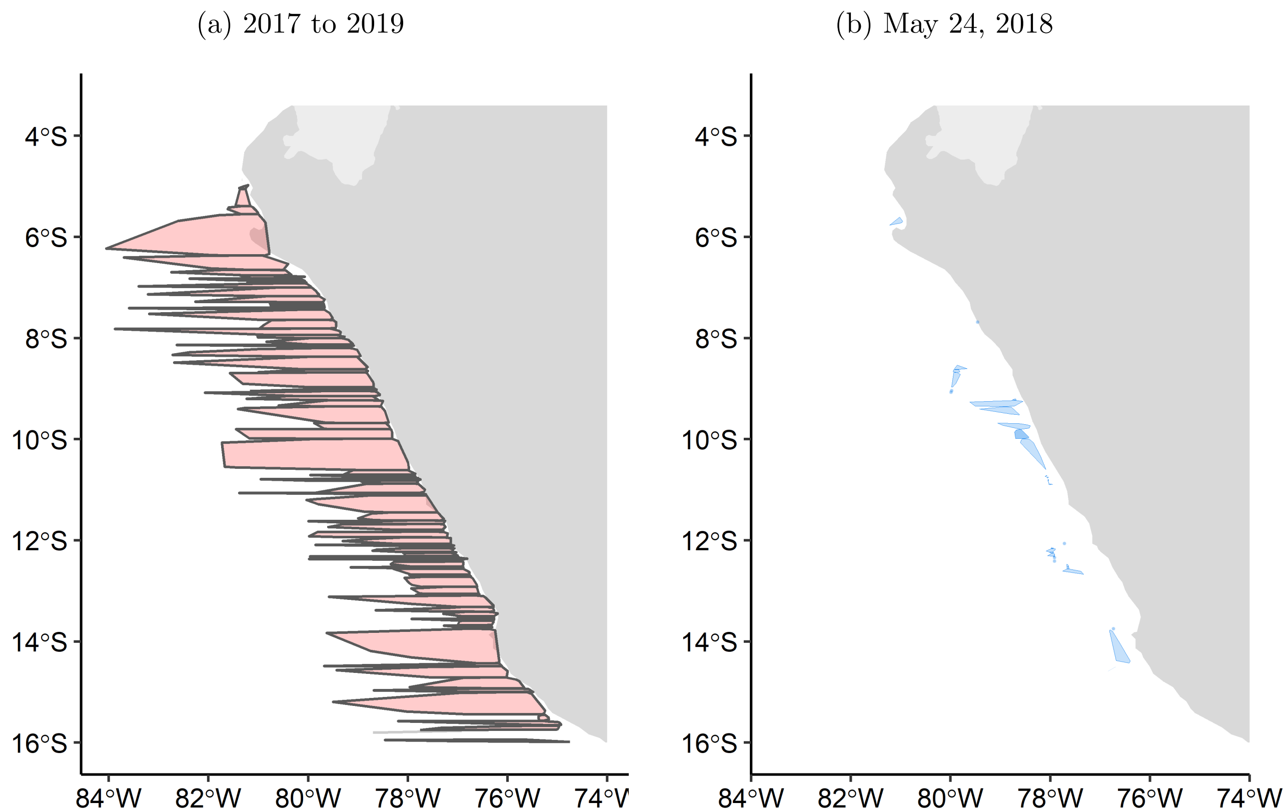
Notes: Each point is a set (a vessel-level fishing operation). The color of each point is the catch per unit effort (CPUE) of that set, which we calculate by adjusting tons caught by vessel and firm characteristics. There are 246,920 sets reported by 806 unique vessels in the electronic logbook data. All vessels are prohibited from fishing within 5 nautical miles (9.3 km) of the coast. Peru is dark grey and Ecuador is light grey.

5.2 Fishing zones

Each set in the electronic logbook data occurs in 1 of 90 fishing zones, which are defined by the regulator and represent distinct, ecologically meaningful fishing grounds. The main decision for vessels, as they relate to day-ahead signals, involves determining which zone to invest hours sailing to. The zones extend far from the coast, but most fishing occurs close to the coast; therefore, the most frequently fished areas within zones are quite small (Figure 2).[17]

[17]Our model assumes that agents choose locations on the line, but the actual fishing choice is two-dimensional. The fact that the coast is long and most fishing occurs close to the coast means that the

Figure 3: Fishing zone polygons



Notes: (a) The regulator defines 90 fishing zones (pink). Our electronic logbook data contains the fishing zone each set occurs in. Here we plot the convex hull of each fishing zone as defined by the sets that occur in that fishing zone during our three years of data. (b) In our analysis, we create zone boundaries based on the sets that occur each day. Here we plot an example of zone boundaries for May 24, 2018.

Figure 3(a) shows the convex hull of each fishing zone as defined by the sets that occur in that zone during our three years of data. The zonal average CPUE ranges from -97.1 to 73.0 and the mean within-zone standard deviation of CPUE is 56.2. Section 6 explains how we identify the ideal fishing ground based on each zone's average CPUE each day. We define zones' polygons each day based on the sets that occurred that day; Figure 3(b) displays an example. The polygon boundaries in Figure 3(a) do not affect our estimate of the distance between a set and the ideal fishing ground.

The zone-level average CPUE uses multiple sets, providing a less noisy measure of the ideal fishing ground, compared to using the single set with the highest CPUE. Since daily data on

model approximates the empirical setting. Section 9 considers the two-dimensional problem.

the actual density of anchoveta do not exist, we use CPUE as its proxy, a standard practice in fishery economics and fishery science (Lynham & Villaseñor-Derbez, 2024; Medoff, Lynham, & Raynor, 2022; Nieto et al., 2017; Zhang & Smith, 2011). Using CPUE as a proxy for anchoveta density assumes that catch conditional on effort is higher in places with greater biomass.

5.3 Public and private information

Firms' public information consists of oceanic geophysical variables measured daily by National Aeronautics and Space Administration (NASA) and National Oceanic and Atmospheric Administration (NOAA) satellites: chlorophyll, sea surface temperature (SST), SST anomaly, sea surface salinity, and sea level anomaly. These variables predict anchoveta abundance (Castillo et al., 2019; Silva et al., 2016).[18] We obtain nine variables in total because NASA and NOAA have multiple satellites (Online Appendix C.1). Satellites obtain different measurements at different times because they have different orbital paths. Moreover, most satellites image only part of the earth each day. By using multiple measurements of SST, for example, we obtain a more complete measurement of daily SST. These data are freely available online. When one of the authors visited two fishing firms in December 2019, both firms monitored these data and communicated them to their fishers.

Our measures of private information are yesterday's CPUE and set locations by all vessels belonging to the same firm. We assume that all vessels within a firm know each other's lagged CPUE, because firms share catch information among their vessels (Section 2). Firms can discourage misreporting in electronic logbook data with two monitoring technologies (Englander, 2023). First, all vessels in the fishery are equipped with tamper-proof transponders that relay vessel location, speed, and course to firms in real-time; this information enables firms to independently verify fishing activities and locations, because fishing appears as vessels moving slowly in a circle (Joo et al., 2015). Second, third-party inspectors measure tons landed at processing plants, which firms can compare to tons reported over the course of a fishing trip. The regulator receives but currently does not publish data from these monitoring technologies.

6 Estimation of r

Here we estimate r, the relative precision of public and private signals about the day's location with the highest stock density.

[18]Dissolved oxygen also predicts anchoveta abundance, but we exclude it from our analysis because it is only available from NOAA at the monthly level. Excluding dissolved oxygen should not meaningfully alter our estimates because it is highly correlated with chlorophyll and sea surface temperature (Kim et al., 2020).

The unobserved state of nature in our analytic model is θ, the location with the highest stock density. The empirical application generalizes the analytic specification by assuming that there might be several "best local zones" instead of a single global ideal zone. The generalization takes into account the fishery's large area and transportation costs, which prevent many boats from traveling to the globally ideal zone in a day.[19] It also recognizes that the stock distribution may be patchy, a feature of many fisheries (Sanchirico & Wilen, 2002). That is, instead of assuming that stock density falls monotonically with the distance to the globally ideal zone, we merely assume that the density falls monotonically with the distance from the best local zone. It is as if the single fishery is broken into several smaller fisheries. The data identifies these smaller fisheries, whose boundaries change daily.

The unit of observation is a set i, performed by a vessel belonging to firm f, and occurring in zone z on day t. The location of a set is k_{ifzt}. We suppose that each set could have occurred in a group of feasible zones, which we define as those within 126 km of the zone where the set actually occurred. We chose this radius because it is the median distance among fishing trips calculated in Joo et al. (2015).[20] For a set in a given zone, we define its best local zone as the zone with the highest average CPUE within 126 km. We denote this best local zone as $\hat{\theta}_{izt}$. The dependent variable in our regressions is $\left(\left\|\left(k_{ifzt}-\hat{\theta}_{izt}\right)^b\right\|\right)^2$, the squared distance between the boundary of the zone in which set i occurs and the boundary of the best local zone. The superscript b indicates that we are taking the distance between two boundaries (not between two points), and $\|\cdot\|$ is the Euclidean distance. Sets in the same zone therefore produce the same value of the dependent variable.

In our empirical model, the variable of interest for a firm is the best local zone. Our regressions measure the public and private signals' ability to predict the deviation between sets and this best local zone. To the extent that public and private signals predict this deviation, they can also provide information about the best location. The variance of this prediction is analogous to the variance of the signal; both measure deviations between the predicted best location and the actual best location. We estimate σ_y^2, the variance of the public signal, with the residual sum of squares (RSS) from the following linear regression:

$$\left(\left\|\left(k_{ifzt}-\hat{\theta}_{izt}\right)^b\right\|\right)^2 = \psi X_{ifzt} + e_{ifzt} \tag{10}$$

[19]The length of the fishery exceeds 1,500 km. At a typical cruising speed of 20 km per hour, it would take more than 3 days to travel along the coast from the fishery's southern boundary to its northern boundary (Peraltilla & Bertrand, 2014).

[20]Joo et al. (2015) use hourly vessel location data to calculate the distance vessels travel between leaving port and landing their catch at processing plants. They also calculate each trip's maximum distance from the coast. The median maximum distance from the coast is 25 km.

where ψ are regression coefficients, X_{ifzt} is a matrix of public signals, e_{ifzt} is the error term, and all other terms are defined above.

The matrix X_{ifzt} includes 16,290 public signal predictor variables derived from the satellite measurements of oceanic geophysical variables described in Section 5.3. For each geophysical variable, we record the nearest measured value to each set, an interpolated value at the set's location, and the distance between the set and the nearest measured value. We compute transformations of these predictor variables and include pairwise interactions between them to capture complex relationships between public signals and the dependent variable (Online Appendix C.2).

To avoid overfitting Equation 10, we randomly assign 75% of the days to the training set, resulting in a training set with 183,055 observations from 321 out of 427 days. The remaining data comprises the test set.[21] We divide the data at the day-level because we want to assess how predictive public information would be of squared distance to the best local zone on a given day if we did not observe the actual best local zone that day.

We estimate Equation 10 with lasso, a type of penalized linear regression. We perform 10-fold cross-validation to choose the optimal penalty term (Online Appendix C.4), which results in a value of 572. Using this optimal penalty term, we fit Equation 10 on the entire training set. The lasso regression retains 24 predictor variables with non-zero coefficients. We use the regression coefficients to predict squared distance to the best zone in the test set.

The RSS in the test set—the squared sum of the difference between the actual and the predicted squared distance to the best local zone—is 9.19956e+11, compared to a total sum of squares for the test set of 1.10446e+12. Our test R^2 is therefore 0.167. Fishing opportunities are highly variable within a day and over small spatial scales, and our public signals are coarse and incomplete (e.g., missing a satellite image for a variable in a given location-day). Our public signals are nonetheless somewhat informative. We could obtain a much higher R^2 if we predicted squared distance to the best local zone in-sample (without evaluating our predictions in a test set), but doing so would overestimate the predictive power of the public signals.

To estimate σ_x^2, the variance of the private signal, we replace X_{ifzt} in Equation 10 with a matrix of private signals and then repeat the same procedure we used to estimate σ_y^2. For each set i by a vessel in firm f on day t, we compute functions of CPUE and fishing locations among sets by vessels in firm f on day $t-1$, such as the CPUE of the nearest set from the previous day. We calculate transformations and include pairwise interactions between all predictor variables, resulting in 153 private signal variables (Online Appendix C.3).

Estimating Equation 10 with our optimal shrinkage penalty (161) retains 8 predictor

[21]This random assignment is irrespective of which fishing season a day occurs in, so the training and test sets are approximately temporally balanced within fishing seasons.

Table 1: Public and private signal parameter estimates

σ_y^2	σ_x^2	r	δ
9.19956e+11	1.07994e+12	0.923	0.540

variables with non-zero coefficients. Our estimate of σ_x^2 is 1.07994e+12, compared to the same total sum of squares as above. The private signals are less informative than the public signals; the private information R^2 in the test set is 0.022.

Our estimate of r is therefore 0.923 ($\sqrt{9.19956e+11}$ divided by $\sqrt{1.07994e+12}$). Table 1 displays our parameter estimates.

Given the two estimated error terms (e_{ifzt}), we can calculate their correlation, ρ, as 0.247. This value exceeds the upper limit required for positive definiteness when $n = 135$ (Section 3). The decision rule is invariant to n in the absence of information clubs or global information sharing, so we do not require that all parameter estimates satisfy a constraint involving n (Section 4).

Using our estimate of r and setting $\rho = 0$ in the formula for δ (Equation 1), we obtain the estimate $\delta = 0.540$; the BLUE of θ assigns 0.54 weight to the public signal and 0.46 weight to the private signal. With $\rho = 0.247$, $\delta = 0.553$. Online Appendix C.5 considers alternative specifications and demonstrates the robustness of our results.

7 Estimation of τ

The parameter τ measures the benefit of dispersion—fewer nearby sets—relative to the cost of locating farther from the ideal fishing ground.

We compute dispersion facing set i as

$$D_{it} = \frac{1}{M_t} \sum_{j \neq i} \left(\tilde{k}_{it} - \tilde{k}_{jt} \right)^2, \tag{11}$$

where $\tilde{k}_{it} - \tilde{k}_{jt} = min(\|k_{it} - k_{jt}\|, 126)$ and M_t is the number of sets on day t. There are 891 sets on the median day and the median set has 312 sets within 126 km. For sets j farther than 126 km from set i, we adjust the distance values to 126 km, since sets farther than 126 km all likely have the same negligible effect on the CPUE of set i; 126 km is the same radius we used in Section 6. We refer to D_{it} as dispersion because it measures the distance between set i and all other sets on that day. We calculate dispersion for each set; sets in the same zone have (slightly) different dispersion values.

To measure the cost of locating farther from the ideal fishing ground, we begin with

Section 6's best local zone specification to estimate the ideal location $\hat{\theta}_{izt}$ for each set i. Conditional on the distribution of other sets, expected catch from set i is higher if it is closer to the ideal location, despite the associated higher density of sets and resulting increase in congestion costs.[22] As a result, aggregate harvest per unit of effort at location i, denoted $CPUE_i$, is higher (in expectation) the closer location i is to the ideal location. We therefore have a proxy for the set's distance from the ideal location given the estimate of the ideal location, $\hat{\theta}_{izt}$, and the location of set i, k_{izt} (which is equivalent to k_{it}, but now includes the z subscript to reflect the role of zones). This proxy equals the dependent variable from Section 6 (Equation 10); here, however, we omit the firm f subscript because it is not relevant for our estimation of τ. We standardize squared distance to the best local zone and dispersion (subtracting their mean and dividing by their standard deviation) so that their regression coefficients are directly comparable to each other.

We estimate the following equation with ordinary least squares regression:

$$CPUE_{izt} = \kappa + BD_{it} + A\left(\left\|\left(k_{izt} - \hat{\theta}_{izt}\right)^b\right\|\right)^2 + e_{izt} \tag{12}$$

where κ, A, and B are regression coefficients and e_{izt} is the error term. We two-way cluster standard errors at the level of date t and zone z. We estimate τ as B divided by $-A$. The negative of A measures the benefit of being closer to the best local zone.

We expect that the effect of dispersion on CPUE is positive ($B > 0$). All else equal, more dispersion should increase CPUE because there are fewer sets near set i (less depletion of the local anchoveta population and less physical congestion from nearby vessels). We expect $A < 0$; as set i occurs farther from the best local zone, CPUE should decrease.

Column 1 of Table 2 displays our estimates of B, A, and τ corresponding to Equation 12. Our estimates of B and A have the expected signs. A 1 standard deviation increase in dispersion increases CPUE by 1.596 tons, while a 1 standard deviation increase in squared distance to the best local zone decreases CPUE by 3.558 tons. We estimate $\tau = 0.448$.

While the negative relationship between CPUE and squared distance to the best local zone is partly mechanical, our estimate of the relationship between CPUE and dispersion could be confounded by not observing the biomass of anchoveta at every location-time. Biomass likely increases CPUE and decreases dispersion: location-times with high biomass will likely have both higher CPUE and lower dispersion (sets clustered close to each other). Our estimate of B (and thus of τ) in Column 1 of Table 2 may therefore be biased downward.

We attempt to alleviate this omitted variable bias by implementing the post-double-

[22] This monotonicity is due to the concavity of the agent's optimization problem: it is better to be where the fish are more abundant.

Table 2: Estimates of τ

	CPUE	
	(1)	(2)
Dispersion	1.596	0.496
	(0.954)	(1.013)
DistToBest^2	-3.558	-3.894
	(0.892)	(0.734)
τ	0.448	0.127
	(0.265)	(0.259)
Controls	No	Yes
Adj. R^2	0.005	0.066
N	246,920	246,920

Notes: The dependent variable is catch per unit effort (CPUE). Mean CPUE is 0 by construction; the standard deviation is 51.8. We estimate τ as the Dispersion coefficient divided by the negative of the DistToBest^2 coefficient. We estimate the standard error of τ with the delta method. Our post-double-selection procedure retains 265 control variables that predict CPUE or Dispersion (Column 2). We two-way cluster standard errors by date and zone.

selection method of Belloni, Chernozhukov, and Hansen (2014). Since omitted variables bias is caused by variables that affect both the dependent variable (CPUE) and the independent variable of interest (dispersion), post-double-selection identifies the predictors of either variable, and controls for these predictors in a third regression of the dependent variable on the independent variable of interest. We include as potential predictors all of the public signal variables that we used to estimate σ_y^2 in Section 6, squared distance to the best local zone, and new potential predictors of biomass and fishing costs (Online Appendix C.6).

Column 2 of Table 2 displays our estimates when we control for the 265 unique predictor variables retained by the two lasso regressions. Instead of becoming larger, which would be consistent with a downward biased coefficient in Column 1, the coefficient on dispersion becomes smaller. This result provides evidence against the concern that omitting biomass from Equation 12 biases the dispersion coefficient downward. Thus, we use the Column 1 estimate of τ in our welfare calculations in Section 8 because that estimate derives from an equation that matches our model more closely (i.e., Equation 2 does not contain control variables). Online Appendix C.6 confirms the robustness of our estimates to alternative specifications of CPUE, dispersion, and travel costs.

8 Welfare effects of information

We use our estimates from Sections 6 and 7 and the theory from Sections 3 and 4 to study the welfare effects of an exogenous increase in a signal's precision, and of the reallocation of information by means of information clubs or global information sharing.

Exogenous increase in the precision of signals. For our point estimates $(r, \tau) = (0.923, 0.448)$, we use Equations 6 and 7 to obtain the elasticity of welfare with respect to the precision of public information as $\alpha = 0.02$, and the elasticity with respect to the precision of private information as $\beta = 1.98$. These values align with Figure 1, which locates our point estimates at X, in the region where greater precision of either signal increases welfare. The elasticity α is positive but close to zero; a small parameter change can flip the sign of this elasticity, but because it would remain small in magnitude, the parameter change would not alter the policy message: more precise private signals substantially raise welfare, but a more precise public signal likely has a small welfare effect.

Information clubs. The decision rule and the payoff in Lemma 1 are differentiable functions of c, but the symmetry assumption means that the domain of c is a subset of integers between 1 and n. For example, with $n = 30$, the symmetric club size is $c \in \{1, 2, 3, 5, 6, 10, 15, 30\}$. For our parameter values, the relevant derivatives are monotonic in c, so there is no loss in generality in ignoring this integer constraint.

Using Equation 3, Online Appendix B.3 shows that

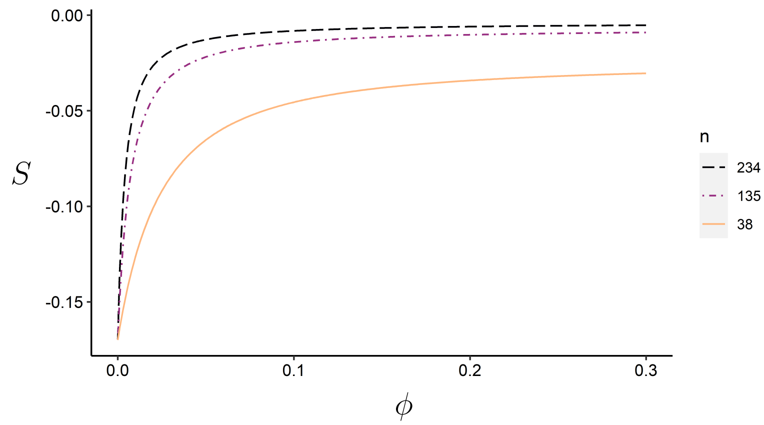
$$\text{sign}\left(\frac{d\gamma^{NE,c}}{dc}\right) = \text{sign}\left[\left(r\left(1 - n\left(1 - \tau\right)\right)\right)\right].$$

With clubs, γ is still the weight on the public signal, but now $1-\gamma$ is the agent's weight on the average of club members' private signals. A larger club makes that average more informative. Therefore, an increase in the club size reduces the equilibrium weight on the public signal.

The function S defined in Equation 4 is a normalized payoff: the per-firm payoff divided by the scaling factor $\frac{A}{2}\sigma_x^2$. To compare scenarios where a club comprises the same fraction of the total industry but n is different, we define $\phi = \frac{c-1}{n-1}$: the number of other firms in i's club as a fraction of the number of other firms. For $\phi = 0$, $c = 1$, and for $\phi = 1$, $c = n$. Using this definition, our point estimates $(r, \tau) = (0.923, 0.448)$, and Equation 4, Equation 37 in Online Appendix B.3 reports the normalized welfare, a function of (ϕ, n).

Figure 4 graphs this normalized payoff as a function of ϕ for $n \in \{38, 135, 234\}$, the number of firms fishing per day at the 25th, 50th, and 75th percentiles of our data. For positive ϕ, an increase in n means that there are more firms outside of i's club, leading to increased dispersion and a higher payoff. This figure also illustrates the fact discussed above, that in the absence of clubs and global information sharing, the payoff per firm does not depend on n.

Figure 4: Normalized payoff as a function of club size



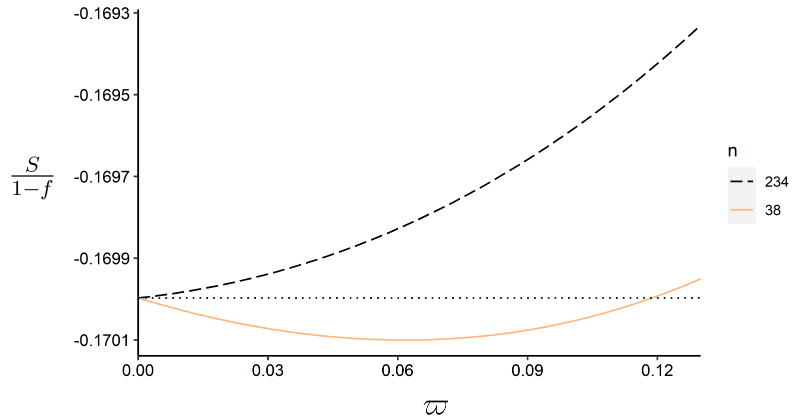
Notes: The normalized payoff S, as a function of ϕ for n equal to the 25th, the 50th, and the 75th percentile of the number of firms fishing per day.

The figure shows a substantial percent increase in payoff as ϕ rises from 0 to 0.15 (where other firms in i's club comprise 15% of the population), and a negligible increase for larger clubs. These results suggest that information clubs would likely increase industry payoffs. But the cost of forming such clubs could limit their creation.

Global information sharing. Using Sections 3 and 4 and the point estimates $(r, \tau) = (0.923, 0.448)$, we find that for $n > 58$, welfare monotonically increases with the amount of information sharing. For smaller n, welfare falls for small ϖ and then rises. For example, at $n = 38$ (the number of firms fishing at the 25th percentile day), information sharing raises welfare only if it reduces the variance of the public signal by 11.9%, i.e., for $\varpi > 0.119$. However, for larger n (including $n = 135$, the median daily number of firms), even a small amount of information sharing raises welfare. This qualitative result is intuitive. The larger is n, the smaller is the fraction (f) of private information that the regulator needs to take from each firm, to achieve a given reduction in the standard deviation of the public signal. In our empirical context, where both public and private information are valuable ($\alpha > 0,\ \beta > 0$), firms are better off when they achieve a more precise public signal at a smaller reduction in the precision of their private information.

To reach these conclusions, we first invert Equation 8 to write f as a function of ϖ. Substi-

Figure 5: Normalized payoff as a function of global information sharing



Notes: The normalized payoff, $\frac{S}{1-f}$, as a function of ϖ for $n = 38$ (the 25th percentile of the number of firms fishing per day) and $n = 234$ (the 75th percentile of the number of firms fishing per day). For $n = 234$, welfare increases with the amount of global information sharing. For $n = 38$, information sharing reduces welfare unless $\varpi > 0.119$ (crossing point indicated by horizontal dotted line).

tuting the result into Equation 9, we compute the elasticities of σ_x and σ_y with respect to ϖ:[23]

$$\begin{aligned} \epsilon_{\sigma_x,\varpi} \equiv \frac{d\sigma_x}{d\varpi}\frac{\varpi}{\sigma_x} = \frac{\varpi}{(1-\varpi)}\left(\frac{1}{2(r^2n(1-\varpi)-\varpi)}\right) \geq 0 \\ \epsilon_{\sigma_y,\varpi} \equiv \frac{d\sigma_y}{d\varpi}\frac{\varpi}{\sigma_y} = \frac{-\varpi}{(1-\varpi)} \leq 0. \end{aligned} \tag{13}$$

Using Equations 6, 7, and 13, we then write the elasticity of the payoff with respect to ϖ:

$$\begin{aligned} \frac{dP}{d\varpi}\frac{\varpi}{|P|} &= \frac{dP}{d\sigma_x}\frac{\sigma_x}{|P|}\frac{d\sigma_x}{d\varpi}\frac{\varpi}{\sigma_x} + \frac{dP}{d\sigma_y}\frac{\sigma_y}{|P|}\frac{d\sigma_y}{d\varpi}\frac{\varpi}{\sigma_y} = (sign(P)2+\alpha)\,\epsilon_{\sigma_x,\varpi} - \alpha\epsilon_{\sigma_y,\varpi} \\ &= \left[(sign(P)2+\alpha)\left(\frac{1}{2(r^2n(1-\varpi)-\varpi)}\right) + \alpha\right]\frac{\varpi}{1-\varpi}. \end{aligned} \tag{14}$$

At our point estimates, $P < 0$, $\alpha = 0.02$ and $r = 0.923$. For ϖ positive but small, the term in the square brackets in the second line of Equation 14 is approximately equal to

$$\frac{-2+0.02}{2r^2n} + 0.02 = \frac{-1.1621+0.02n}{n}.$$

This result shows that a small amount of information sharing ($\varpi \approx 0, \varpi > 0$) increases wel-

[23]Equation 8 and $f \leq 1$ imply that $\varpi \leq \frac{nr^2}{nr^2+1}$, leading to the inequality in the first line of Equation 13. The elasticities are for the post-transfer moments, σ'_x, σ'_y; we write σ_x, σ_y in order to simplify the notation.

fare if $n > 58$ and lowers welfare for smaller n. Figure 5 illustrates this result, showing the graph of normalized payoff as a function of ϖ, for $n = 38$ and $n = 234$.

9 Extensions

Here we consider three extensions: allowing firms to choose a location in two-dimensional space instead of on the real line; a dynamic model that accommodates stock externalities; and games in which the decision variable is emissions rather than location.

Actions in two-dimensional space. This extension requires that firms receive signals of both the latitude and the longitude of the (local) best zone. If firms receive both a public and a private signal for each of these, then they receive four signals, and have to make two decisions, their latitude and longitude. Therefore, a firm would condition both of its decisions on all four signals. Because the weights for each decision sum to one, the firm's linear decision rule requires three weights for each decision, a total of six coefficients.

When all agents use linear decision rules, a firm's payoff is a quadratic function of the weights in its decision rules. We now have six first order conditions, which are linear functions of the weights that other firms use. With symmetry (because all firms use the same rule in equilibrium) we obtain a system of six equations that (in general) can be solved numerically to obtain the coefficients of the equilibrium linear decision rules as functions of the signals' moments and of the parameter τ. The six-dimensional equilibrium problem would be too complicated to yield analytic insight, but it could be studied using parameter estimates.

Stock externalities. For reasons discussed in Section 2, we doubt that stock externalities are of first order importance in the anchoveta fishery, but they likely are in other resource settings. There has been an explosion of the literature on estimating dynamic games, particularly in the field of industrial organization (Aguirregabiria, Collard-Wexler, & Ryan, 2021); perhaps these methods can be adapted to study a dynamic version of our problem.

The components of such a model would include a flow payoff and two sources of dynamics. Ignoring growth within a season, we need one stock variable to keep track of cumulative harvest within the season, to account for the possibility that fish become scarcer and harvest costs higher as the season progresses. Second, each firm needs to keep track of its own seasonal harvest to determine when it reaches its quota. The number of boats whose quota constraint is slack would weakly decrease with the season. The flow payoff could be similar to the payoff in our static model, except that it would also depend on cumulative harvest to account for stock-related harvest costs. The details of any such model would have to be guided by the empirical setting.

Emissions games. The game in which firms choose emissions closely resembles those in which firms choose price or quantity. Suppose that firms receive signals about an unknown technology parameter that affects their abatement costs; these cost parameters might be correlated across firms. A regulator fixes a supply schedule or a number (in the case of standard cap and trade) for emissions permits. Permits are auctioned or otherwise distributed, and a market for permits opens. Vives (2014) discusses a version of this problem in which firms receive only a private signal. Cantillon and Slechten (2018) study a multiperiod emissions game, focusing on the ability of markets to aggregate information. There is scope for further research in this area, including consideration of both public and private signals about technology; the increased use of emissions markets may produce opportunities for empirical work.

10 Conclusion

A rich theoretical literature recognizes the differing equilibrium effects of more precise public and private information, and the possibility that either might lower welfare by changing equilibrium behavior. A rich empirical literature estimates the value of information under the assumption that if information is not useful, agents can ignore it without cost: information always has (weakly) positive value (Duflo et al., 2018). Thus, the theory on the ambiguous value of information has rarely been tested empirically; the empirical literature on the value of information has generally ignored the insights from theory. Our results show the value of combining a micro-founded model, careful econometrics, and numerical analysis.

We obtain a full analytic characterization of the welfare effects of more precise information over a larger parameter space than in previous papers. Our model contains two primary parameters: the relative precision of private versus public information, and the importance of congestion relative to proximity to the ideal location. We estimate these parameters using high-resolution data from the world's largest fishery by catch volume, the Peruvian anchoveta fishery. We find that private signals are about 8% less precise than public signals and congestion is about 45% as important to profits as is fishing close to the ideal fishing ground. Accordingly, improving the precision of private information would increase welfare, but improving the precision of public information would have a negligible effect on welfare.

Our model nests two kinds of information sharing. We calculate that information clubs, in which members share private signals, would increase welfare. Even a small amount of global information sharing would also typically increase welfare because the number of active firms is large; the regulator requires less private information from each firm in order to achieve a given increase in the precision of the public signal in this industry.

The primary policy implication is that anchoveta firms would benefit from more precise private information about the location of fish stocks. Firms could increase the precision of their private information by investing in buoy-based acoustic technology that continuously transmits real-time estimates of anchoveta biomass (Brehmer et al., 2019; Simmonds et al., 2009). Second, the regulator could implement global information sharing by publishing near real-time catch data (Englander, 2023).

These conclusions are in line with theory's recognition that, in the presence of congestion (or a similar negative externality) more precise public information is less likely to be beneficial, compared to more precise private information. Public information encourages agents to act in the same manner, e.g., all move to the same fishing ground, thereby exacerbating the externality. The particularity of our results (to the context of the Peruvian anchoveta fishery) cautions against attempts to generalize them to other fisheries, or other settings where externalities are important. This particularity also shows the value of combining theory with empirical tools. This combination enables us to address a type of policy question not accessible to an atheoretical approach.

A Appendix: Technical information

We first discuss the relation between our model and AP07 and then sketch the proofs and some derivations. These are computation-intensive, but do not provide additional insight. Therefore, we relegate this material to Online Appendix B, and here we provide an overview. Online Appendix C provides more information regarding the data and empirical analysis.

AP07's footnote 5 notes that Morris and Shin's (2002) uniqueness proof can be adapted to show that the unique Bayesian Nash equilibrium in this game is linear in strategies, provided that (in our notation) $\tau < 0.5$. This restriction is innocuous in AP07, because they study both the noncooperative Bayesian Nash equilibrium and also a "team problem" in which firms choose location to maximize collective welfare, but without sharing information. Their sufficient condition for concavity of payoffs in the team problem implies $\tau < 0.5$. We are interested only in the noncooperative setting, where the weaker restriction $\tau < 1$ is necessary and sufficient for concavity of the agent's problem. We therefore consider the entire parameter region $\tau < 1$; for $0.5 \leq \tau < 1$, where AP07's sufficient condition for uniqueness is not satisfied, we assume that the equilibrium is linear in information. We treat firms as choosing the weight γ in their decision rule, whereas AP07 have firms choose k_i. Comparing our Equation 3 evaluated at $c = 1$ with their Equation 8 on page 1112 provides a consistency check, confirming that the two approaches are equivalent.

Online Appendix B.1 proves Lemma 1 and the two propositions. In writing an agent's objective and first order condition when $c > 1$, we recognize that the agent knows the information other club members have, and therefore can predict their actions. In contrast, the agent takes expectations of the actions of members of different clubs.

To prove the Lemma, we substitute the public signal weights γ (the weight that agent i intends to use) and η (the weight that this agent believes other agents will use) into i's objective. We then replace the signals with their definitions (e.g., $y = \theta + \varepsilon_y$) and cancel the θs. Taking expectations with respect to the noise $(\varepsilon_y, \{\varepsilon_{x_i}\})$ produces i's expected welfare as a function of γ and η and the signals' moments. Evaluating the first order condition with respect to γ at a symmetric equilibrium ($\gamma = \eta$) produces Equation 3. Evaluating the payoff at this symmetric equilibrium produces Equation 4.

To prove the two propositions, we set $c = 1$ and $\rho = 0$ and rewrite the expression for the equilibrium payoff in terms of (σ_y, σ_x). Taking the derivative of the payoff with respect to these two moments and simplifying completes most of the proofs. However, to establish Proposition 1.iii we use L'Hopital's Rule to evaluate the payoff at the limiting values of σ_y.

Online Appendix B.2.2 uses Lemma 1 with $c = 1$ to construct (for $\rho \lesseqgtr 0$) the boundaries of the sets in the (τ, δ) plane where the derivatives of the payoff with respect to σ_y and σ_x do not change signs. Figure 1 shows the boundaries for $\rho = 0$. Figures B1 and B2 show the boundaries for public and private information with $\rho > 0$. The construction of these boundaries is somewhat intricate because the welfare derivatives are ratios of polynomials. To avoid a complicated taxonomy, we restrict $\tau < 1 - \rho^2$ to ensure that the denominator of these ratios does not pass through a zero (at which point the derivative is undefined). Online Appendix B.2.3 computes the elasticities of the payoff with respect to the precision of public and private information, shown in Equations 6 and 7.

Online Appendix B.3 uses Lemma 1 to take the derivative of the equilibrium weight ($\gamma^{NE,c}$) with respect to the club size. To construct Figure 4, we use $c = 1 + (n-1)\phi$ to eliminate c from the expression for S, writing the normalized payoff as a function of the point estimates of the model primitives and (n, ϕ).

Online Appendix B.4 describes the model in which each agent receives m private signals. This generalization enables us to consider the possibility that the regulator collects only a fraction of agents' private information. We choose the moments of these m signals so that, absent information sharing, the model is equivalent to our original model in which each agent receives a single private signal. We then compute the moments of the enhanced public signal and the degraded private signal, when the regulator collects the fraction f of each firm's private information. In the process, we derive Equations 8 and 9. The appendix considers the general case $\rho \lesseqgtr 0$, so it also shows how global information sharing affects this correlation.

Data Availability Statement

The data and code underlying this research are available on Zenodo at https://doi.org/10.5281/zenodo.15122098.

References

Aguirregabiria, V., Collard-Wexler, S., & Ryan, S. (2021). Dynamic games in empirical insustrial organization. *Handbook of Industrial Organization*, *4*, 225–343.

Angeletos, G.-M., & Pavan, A. (2007). Efficient use of information and social value of information. *Econometrica*, *75*, 1103–1142.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, *81*(2), 608–650.

Bergemann, D., & Morris, S. (2013). Robust predictions in games with incomplete information. *Econometrica*, *81*(4), 1251–1308.

Bolton, P., & Harris, C. (1999). Strategic experimentation. *Econometrica*, *67*(2), 349–374.

Brehmer, P., Sancho, G., Trygonis, V., Itano, D., Dalen, J., Fuchs, A., Faraj, A., & Taquet, M. (2019). Towards an autonomous pelagic observatory: experiences from monitoring fish communities around drifting FADs. *Thalassas: An International Journal of Marine Sciences*, *35*, 177–189.

Brown, G. (1974). An optimal program for managing common property resources with congestion externalities. *Journal of Political Economy*, *82*, 163–173.

Callander, S., & Harstad, B. (2015). Experimentation in federal systems. *The Quarterly Journal of Economics*, *130*(2), 951–1002.

Calvó-Armengol, A., De Martí, J., & Prat, A. (2015). Communication and influence. *Theoretical Economics*, *10*(2), 649–690.

Cantillon, E., & Slechten, A. (2018). Information aggregation in emissions markets with abatement. *Annals of Economics and Statistics*, (132), 53–79.

Castillo, R., Rosa, L. D., Diaz, W. G., Madureira, L., Gutierrez, M., Vásquez, L., & Koppelmann, R. (2019). Anchovy distribution off Peru in relation to abiotic parameters: A 32-year time series from 1985 to 2017. *Fisheries Oceanography*, *28*(4), 389–401.

Clarke, R. (1983). Collution and the incentives for information sharing. *Bell Journal of Economics*, *14*, 383–394.

Duflo, E., Greenstone, M., Pande, R., & Ryan, N. (2018). The value of regulatory discretion: Estimates from environmental inspections in India. *Econometrica*, *86*(6), 2123–2160.

Englander, G. (2023). Information and Spillovers from Targeting Policy in Peru's Anchoveta Fishery. *American Economic Journal: Economic Policy*, *15*(4), 390–427.

FAO. (2018). The State of World Fisheries and Aquaculture 2018 - Meeting the Sustainable Development Goals [Food and Agricultural Organization of the United Nations (FAO)].

Fréon, P., Sueiro, J. C., Iriarte, F., Evar, O. F. M., Landa, Y., Mittaine, J.-F., & Bouchon, M. (2014). Harvesting for food versus feed: A review of Peruvian fisheries in a global context. *Reviews in Fish Biology and Fisheries*, *24*(1), 381–398.

Gal-or, E. (1986). Information transmission – Couornot and Bertrand equilibria. *Review of Economic Studies*, *53*(2), 85–92.
Gatewood, J. (1984). Cooperation, competition, and synergy: Information-sharing groups among southeast Alaskan salmon seiners. *American Ethnologist*, *11*, 350–370.
Hansman, C., Hjort, J., León-Ciliotta, G., & Teachout, M. (2020). Vertical Integration, Supplier Behavior, and Quality Upgrading among Exporters. *Journal of Political Economy*, *128*(9), 3570–3625.
Huang, L., & Smith, M. (2014). The dynamic efficiency costs of common-pool resource exploitation. *American Economic Review*, *104*, 4071–4103.
IMARPE. (2017). *Situación del Stock Norte-Centro de la anchoveta peruana (Engraulis ringens) al 01 de Noviembe de 2017 y perspectivas de explotación para la segunda temporada de pesca de 2017*. Instituto del Mar del Perú (IMARPE).
Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *The Quarterly Journal of Economics*, *122*(3), 879–924.
Joo, R., Salcedo, O., Gutierrez, M., Fablet, R., & Bertrand, S. (2015). Defining fishing spatial strategies from VMS data: Insights from the world's largest monospecific fishery. *Fisheries Research*, *164*, 223–230.
Kim, Y. H., Son, S., Kim, H.-C., Kim, B., Park, Y.-G., Nam, J., & Ryu, J. (2020). Application of satellite remote sensing in monitoring dissolved oxygen variabilities: A case study for coastal waters in Korea. *Environment International*, *134*, 105301.
Lynham, J., & Villaseñor-Derbez, J. C. (2024). Evidence of spillover benefits from large-scale marine protected areas to purse seine fisheries. *Science*, *386*(6727), 1276–1281.
Marcoul, P. (2020). Fish abundance information and fishery's aggregate behavior. *Journal of the Association of Environmental and Resource Economists*, *7*(5), 933–964.
Medoff, S., Lynham, J., & Raynor, J. (2022). Spillover benefits from the world's largest fully protected MPA. *Science*, *378*(6617), 313–316.
Morris, S., & Shin, H. S. (2002). Social value of public information. *American Economic Review*, *92*(5), 1521–1534.
Myatt, D. P., & Wallace, C. (2012). Endogenous Information Acquisition in Coordination Games. *The Review of Economic Studies*, *79*(1), 340–374.
National Aeronautics and Space Administration. (2024). PACE Mission: Plankton, Aerosol, Cloud, ocean Ecosystem [Accessed: 2024-12-10]. https://pace.gsfc.nasa.gov/
Natividad, G. (2016). Quotas, productivity, and prices: The case of anchovy fishing. *Journal of Economics & Management Strategy*, *25*(1), 220–257.
Nieto, K., Xu, Y., Teo, S. L., McClatchie, S., & Holmes, J. (2017). How important are coastal fronts to albacore tuna (Thunnus alalunga) habitat in the Northeast Pacific Ocean? *Progress in Oceanography*, *150*, 62–71.
Peraltilla, S., & Bertrand, S. (2014). In situ measurements of the speed of Peruvian anchovy schools. *Fisheries Research*, *149*, 92–94.
PRODUCE. (2018). Anuario Estadístico Pesquero y Acuícola 2017 [Ministerio de la Producción (PRODUCE)].
PRODUCE. (2020a). Bitácora Electrónica, 2017 to 2019 [DATASET] [Ministerio de la Producción (PRODUCE). Accessed March 9, 2020. Available at https://doi.org/10.5281/zenodo.15122098].

PRODUCE. (2020b). Embarcaciones Pesqueras [DATASET] [Ministerio de la Producción (PRODUCE). Accessed March 14, 2020. Available at https://consultasenlinea.produce.gob.pe/ConsultasEnLinea/consultas.web/embarcacion].

PRODUCE. (2020c). Landings, 2016 [DATASET] [Ministerio de la Producción (PRODUCE). Accessed March 9, 2020. Available at https://doi.org/10.5281/zenodo.15122098].

Sanchirico, J., & Wilen, J. (2002). Bioeconomics of spatial exploitation in a patchy environnriment. *Fisheries economics* (pp. 177–198). Routledge.

Sanz, N., & Diop, S. (2021). Endogenous catch per unit effort and congestion externalities between vessels in a search-matching model: Evidence from the French Guiana shrimp fishery. *Bulletin of Economic Research*, *11*, 1–16.

Silva, C., Andrade, I., Yáñez, E., Hormazabal, S., Barbieri, M. Á., Aranis, A., & Böhm, G. (2016). Predicting habitat suitability and geographic distribution of anchovy (Engraulis ringens) due to climate change in the coastal areas off Chile. *Progress in Oceanography*, *146*, 159–174.

Simmonds, E. J., Gutiérrez, M., Chipollini, A., Gerlotto, F., Woillez, M., & Bertrand, A. (2009). Optimizing the design of acoustic surveys of Peruvian anchoveta. *ICES Journal of Marine Science*, *66*(6), 1341–1348.

SUPNEP. (2017). Convenio Colectivo por Rama de Actividad, SUPNEP - AANEP, 12 de marzo 2017 al 11 marzo del 2022 [Sindicato Univo de Pescadores de Nuevas Embarcaciones del Peru (SUPNEP). http://www.supnep.pe/assets/convenio_colectivo_supnep_aanep_2017_2022.pdf].

Ui, T., & Yoshizawa, Y. (2015). Characterizing social value of information. *Journal of Economic Theory*, *158*, 507–535.

Vives, X. (1984). Duopoly information equilibrium: Cournot and Bertrand. *Journal of Economic Theory*, *34*, 71–94.

Vives, X. (1990). Trade association disclosure rules, incentives to share information, and welfare. *RAND Journal of Economics*, *21*(3), 409–430.

Vives, X. (2014). On the possibility of informationally efficient markets. *Journal of the European Economic Association*, *12*(5), 1200–1239.

Welch, H., Clavelle, T., White, T. D., Cimino, M. A., Van Osdel, J., Hochberg, T., Kroodsma, D., & Hazen, E. L. (2022). Hot spots of unseen fishing vessels. *Science Advances*, *8*(44), eabq2109.

Zhang, J., & Smith, M. (2011). Estimation of a generalized fishery model. *Review of Economics and Statistics*, *93*, 690–699.

Online Appendix for "The value of information in a congested fishery"

Appendix A provides a sketch of the proofs and of derivations and it also serves as a road map to Online Appendix B, which contains the detailed proofs and derivations, and supporting arguments for some of the claims made in the paper. Online Appendix C contains supporting information for the empirical results.

B Online Appendix: Proofs and derivations

We first explain the necessary and sufficient condition for a positive definite covariance matrix of the entire set of public and private signals, $n\rho^2 < 1$. Using the definition $r \equiv \frac{\sigma_y}{\sigma_x}$, the covariance matrix for the public signal and the n private signals can be written as $\sigma_x^2\Gamma$, where the $(n+1) \times (n+1)$ matrix Γ has r^2 in the $(1,1)$ entry, 1's in other diagonal entries, $r\rho$ on the first row and column, and zeros elsewhere. For example, for $n = 3$

$$\Gamma = \begin{pmatrix} r^2 & r\rho & r\rho & r\rho \\ r\rho & 1 & 0 & 0 \\ r\rho & 0 & 1 & 0 \\ r\rho & 0 & 0 & 1 \end{pmatrix}.$$

It can be shown, e.g., using an inductive proof, that the determinant is $|\Gamma| = r^2(1 - n\rho^2)$, i.e., $|\Gamma| > 0 \iff 1 > n\rho^2$. With $1 > n\rho^2$, the principal minors of Γ are also positive.

B.1 Proofs

Proof. (Lemma 1) The definitions in the text imply

$$\tilde{\varepsilon}_s \sim \left(0, \frac{\sigma_x^2}{c}\right), \mathbb{E}\tilde{\varepsilon}_s\varepsilon_y = \rho\sigma_x\sigma_y, \text{ and } \mathbb{E}\tilde{\varepsilon}_s\tilde{\varepsilon}_k = 0 \text{ for } k \neq s. \tag{15}$$

Let agent i be a member of a particular club, indexed by $\tilde{s}_{(i)}$. In a linear equilibrium, agent i uses the decision rule $k_i = \gamma y + (1-\gamma)\,\tilde{x}_{\tilde{s}_{(i)}}$ and agent $j \neq i$ (including agents in the same club as i) uses the rule $k_j = \eta y + (1-\eta)\,\tilde{x}_{s(j)}$, where $\tilde{x}_{s(j)}$ denotes the collective semi-private signal of the club to which agent j belongs. Substituting these linear decision rules into firm

i's payoff, $J(k_i|\{k_j\},\theta)$, gives the payoff as a function of the weights γ and η:

$$\begin{gathered}\tfrac{1}{n-1}\tfrac{B}{2}\left[(c-1)\left(\gamma y+(1-\gamma)\,\tilde{x}_{\tilde{s}}-\left(\eta y+(1-\eta)\,\tilde{x}_{\tilde{s}}\right)\right)^2+\right.\\ \left.\textstyle\sum_{s\neq\tilde{s}} c\left(\left(\gamma y+(1-\gamma)\,\tilde{x}_{\tilde{s}}\right)-\left(\eta y+(1-\eta)\,\tilde{x}_{s}\right)\right)^2\right]\\ -\tfrac{A}{2}\left(\gamma y+(1-\gamma)\,\tilde{x}_{\tilde{s}}-\theta\right)^2.\end{gathered}$$

The term in the first line, multiplied by $c-1$, is the contribution to dispersion provided by the other $c-1$ members of i's club. In a symmetric equilibrium, this contribution is zero: all members of the club have the same information, so they locate in the same position. The second line measures the dispersion provided by members of other $T-1$ clubs, each of which has c members. The third line is the cost of missing the target. The expectation of this payoff is

$$\begin{gathered}P^{club}\equiv\tfrac{1}{n-1}\tfrac{B}{2}\left[(c-1)\left(\gamma y+(1-\gamma)\,\tilde{x}_{\tilde{s}}-\left(\eta y+(1-\eta)\,\tilde{x}_{\tilde{s}}\right)\right)^2+\right.\\ \left.\mathbb{E}\textstyle\sum_{s\neq\tilde{s}} c\left(\left(\gamma y+(1-\gamma)\,\tilde{x}_{\tilde{s}}\right)-\left(\eta y+(1-\eta)\,\tilde{x}_{s}\right)\right)^2\right]\\ -\tfrac{A}{2}\mathbb{E}\left(\gamma y+(1-\gamma)\,\tilde{x}_{\tilde{s}}-\theta\right)^2.\end{gathered}\tag{16}$$

Terms in the first line are observed, so there is no expectations operator there. The first expectations operator, appearing in the second line, is over other clubs' private signals. There are c members of each of these other clubs, and all members of a club have the same information and make the same decision. The second expectations operator, appearing in the third line, is over θ. We simplify the second line of Equation 16 by replacing the signals by their definitions, i.e., $y=\theta+\varepsilon_y$ and $\tilde{x}_{\tilde{s}}=\theta+\tilde{\varepsilon}_{x_{\tilde{s}}}$; we then cancel the θs and then take expectations with respect of ε_y and $\{\tilde{\varepsilon}_{x_{\tilde{s}}}\}$. These steps produce:

$$\begin{gathered}\mathbb{E}\textstyle\sum_{s\neq\tilde{s}} c\left(\left(\gamma y+(1-\gamma)\,\tilde{x}_{\tilde{s}}\right)-\left(\eta y+(1-\eta)\,\tilde{x}_{s}\right)\right)^2=\\ \mathbb{E}\textstyle\sum_{s\neq\tilde{s}} c\left(\gamma\left(\theta+\varepsilon_y\right)+(1-\gamma)\left(\theta+\tilde{\varepsilon}_{x_{\tilde{s}}}\right)-\left(\eta\left(\theta+\varepsilon_y\right)+(1-\eta)\left(\theta+\tilde{\varepsilon}_{x_s}\right)\right)\right)^2=\\ \mathbb{E}\textstyle\sum_{s\neq\tilde{s}} c\left(\gamma\varepsilon_y+(1-\gamma)\,\tilde{\varepsilon}_{x_{\tilde{s}}}-\left(\eta\varepsilon_y+(1-\eta)\,\tilde{\varepsilon}_{x_s}\right)\right)^2=\\ \mathbb{E}\textstyle\sum_{s\neq\tilde{s}} c\left((\gamma-\eta)\,\varepsilon_y+(1-\gamma)\,\tilde{\varepsilon}_{x_{\tilde{s}}}-(1-\eta)\,\tilde{\varepsilon}_{x_s}\right)^2=\\ (T-1)\,c\left[(\gamma-\eta)^2\sigma_y^2+\left((1-\gamma)^2+(1-\eta)^2\right)\tfrac{\sigma_x^2}{c}-2\,(\gamma-\eta)^2\rho\sigma_x\sigma_y\right]=\\ (T-1)\,c\sigma_x^2\left[(\gamma-\eta)^2r^2+(1-\gamma)^2\tfrac{1}{c}+(1-\eta)^2\tfrac{1}{c}-2r\rho\,(\gamma-\eta)^2\right].\end{gathered}\tag{17}$$

Before this sequence of equalities we explained the steps used to obtain the first three equalities in the sequence. To obtain the fourth equality of Sequence 17 we simplify the expectation of the cross product term using

$$\mathbb{E}2\left(\gamma-\eta\right)\varepsilon_{y}\left(\left(1-\gamma\right)\tilde{\varepsilon}_{x_{\tilde{s}}}-\left(1-\eta\right)\tilde{\varepsilon}_{x_{s}}\right)=2\left(\gamma-\eta\right)\left(\left(1-\gamma\right)-\left(1-\eta\right)\right)\rho\sigma_{x}\sigma_{y}=$$
$$-2\left(\gamma-\eta\right)^{2}\rho\sigma_{x}\sigma_{y}.$$

This expression appears as the last term on the fifth line of Sequence 17. We obtain the final line by factoring out σ_x^2 and using the definition of r.

The expected cost of missing of the target is

$$\begin{aligned}&\tfrac{A}{2}\mathbb{E}\left(\gamma y+\left(1-\gamma\right)\tilde{x}_{\tilde{s}}-\theta\right)^{2}=\tfrac{A}{2}\mathbb{E}\left(\gamma\varepsilon_{y}+\left(1-\gamma\right)\tilde{\varepsilon}_{\tilde{s}}\right)^{2}=\\&\tfrac{A}{2}\left(\gamma^{2}\sigma_{y}^{2}+\left(1-\gamma\right)^{2}\tfrac{\sigma_{x}^{2}}{c}+2\gamma\left(1-\gamma\right)\rho\sigma_{x}\sigma_{y}\right)=\\&=\tfrac{A}{2}\sigma_{x}^{2}\left(\gamma^{2}r^{2}+\left(1-\gamma\right)^{2}\tfrac{1}{c}+2\gamma\left(1-\gamma\right)\rho r\right).\end{aligned}\tag{18}$$

Once again, the first expectations operator is over θ; we obtain the first equality in this sequence by replacing the signals with their definitions and then canceling θ. Thus, the second expectations operator is over the shocks $\varepsilon_y, \tilde{\varepsilon}_{x_{\tilde{s}}}$. The second equality follows from taking expectations; the third equality follows from simplification, using the definition of r.

Using Equations 16, 17 and 18 we write agent i's expected payoff

$$\begin{aligned}&P^{club}\equiv\\&\tfrac{c-1}{n-1}\tfrac{B}{2}\left(\gamma y+\left(1-\gamma\right)\tilde{x}_{\tilde{s}}-\left(\eta y+\left(1-\eta\right)\tilde{x}_{\tilde{s}}\right)\right)^{2}\\&+\tfrac{(n-c)}{n-1}\tfrac{B}{2}\sigma_{x}^{2}\left[\left(\gamma-\eta\right)^{2}r^{2}+\left(1-\gamma\right)^{2}\tfrac{1}{c}+\left(1-\eta\right)^{2}\tfrac{1}{c}-2r\rho\left(\gamma-\eta\right)^{2}\right]\\&-\tfrac{A}{2}\sigma_{x}^{2}\left(\gamma^{2}r^{2}+\left(1-\gamma\right)^{2}\tfrac{1}{c}+2\gamma\left(1-\gamma\right)\rho r\right).\end{aligned}\tag{19}$$

The second line uses $n = Tc$ to write $\frac{T-1}{n-1}c = \frac{n-c}{n-1}$.

The derivative of the *first line* after the identity symbol in Equation 19, with respect to γ, is

$$\frac{c-1}{n-1}\frac{B}{2}2\left(\gamma y+\left(1-\gamma\right)\tilde{x}_{\tilde{s}}-\left(\eta y+\left(1-\eta\right)\tilde{x}_{\tilde{s}}\right)\right)\left(y-\tilde{x}_{\tilde{s}}\right).$$

Evaluated at a symmetric equilibrium ($\gamma = \eta$) this expression equals zero. Therefore, the possibility of dispersion between an agent and other club members does not affect either the payoff or the first order condition, and thus does not affect the decision rule. We use this fact to write the payoff more simply by dropping the first line after the identity in Equation 19. Once we drop this term, it is apparent that the club causes the "effective payoff weight" on the dispersion term to fall from B to $\frac{n-c}{n-1}B$.

We also use the definition $\Upsilon \equiv \frac{(n-c)}{n-1}\frac{B}{A} = \frac{(n-c)}{n-1}\tau \leq \tau < 1$; the strict inequality follows from

Assumption 1(ii). The resulting simplified payoff is

$$\begin{gathered}P^{club} \equiv \\ A\sigma_x^2\left[\frac{\Upsilon}{2}\left(\left(\gamma-\eta\right)^2 r^2+\left(1-\gamma\right)^2\frac{1}{c}+\left(1-\eta\right)^2\frac{1}{c}-2r\rho\left(\gamma-\eta\right)^2\right)\right. \\ \left.-\frac{1}{2}\left(\gamma^2 r^2+\left(1-\gamma\right)^2\frac{1}{c}+2\gamma\left(1-\gamma\right)\rho r\right)\right].\end{gathered} \tag{20}$$

We denote the function in square brackets as $\tilde{P}$, so $P^{club} = A\sigma_x^2\tilde{P}$.

The first order condition for maximizing this payoff with respect to γ, $\frac{d\tilde{P}}{d\gamma}=0$, implies

$$cr\Upsilon\left(r-2\rho\right)\eta+\left(cr^2-2cr\rho+1\right)\left(1-\Upsilon\right)\gamma+\left(\Upsilon+cr\rho-1\right)=0. \tag{21}$$

Evaluating this first order condition at the symmetric equilibrium, $\eta=\gamma$, and solving for γ gives the equilibrium weight

$$\gamma^{NE,c}=\frac{1-\Upsilon-cr\rho}{1-\Upsilon+cr^2-2cr\rho},$$

which produces Equation 3. We obtain Equation 4 using Equation 20 (factoring the $\frac{1}{2}$) evaluated at the equilibrium, where $\eta=\gamma$.

The second order condition for the agent's problem uses

$$\frac{d^2\tilde{P}}{d\gamma^2}=\frac{1}{c}\left(\Upsilon-1\right)\left(cr^2-2c\rho r+1\right).$$

We now show that $cr^2-2cr\rho+1>0$. This inequality follows from inspection for $\rho\le 0$, so we need only consider the case $\rho>0$. For this case $cr^2-2cr\rho+1>f(c,r)\equiv cr^2-\frac{2cr}{\sqrt{n}}+1$ (because $\rho<\frac{1}{\sqrt{n}}$), so we need to show that $f(c,r)\ge 0$. Again, we have two cases. For $r\ge\frac{2}{\sqrt{n}}$, $f(c,r)\ge 0$ by inspection (because $c>0$). For the alternative, $r<\frac{2}{\sqrt{n}}$, use the fact that $f(c,r)$ is a quadratic in r with a single root at $r^*=\frac{1}{\sqrt{n}}$ (that is, $f(c,r^*)=0$). The quadratic is positive at $r=0$, and because the unique root is r^* the quadratic is positive for $f\ne r^*$. Thus, the firm's second order condition is satisfied if and only if $\Upsilon<1$. This inequality holds for $c=1$ if and only if $\tau<1$. With this inequality, the second order condition holds for all $c\le n$. □

Proof. (Proposition 1) Set $\rho=0$ and $c=1$ and evaluate the equilibrium payoff, Equation 4, at the equilibrium decision rule, Equation 3, to write the payoff in terms of model primitives, denoted Z:

$$Z\equiv-\tfrac{1}{2}\sigma_y^2 A\sigma_x^2\frac{\sigma_x^2\left(\tau^2-2\tau+1\right)+\sigma_y^2(1-2\tau)}{\left(-\tau\sigma_x^2+\sigma_y^2+\sigma_x^2\right)^2}. \tag{22}$$

Part i. The derivative of Z with respect to σ_y^2 is

$$\begin{aligned} \frac{dZ}{d\sigma_y^2} &= \left(\frac{\frac{1}{2} A\sigma_x^4 (1-\tau)}{\left((\tau-1)\sigma_x^2 - \sigma_y^2\right)^3} \right) \left((1-3\tau)\,\sigma_y^2 + \sigma_x^2 \left(\tau^2 - 2\tau + 1\right) \right) \\ &= \tfrac{1}{2} A\delta^2 \frac{1-\tau}{(1-\delta\tau)^3} \left(3\tau - \tau\delta - \tau^2\delta - 1\right). \end{aligned} \tag{23}$$

Because $\tau < 1$ and $\delta \leq 1$, the coefficient, $\frac{1}{2}A\delta^2 \frac{1-\tau}{(1-\delta\tau)^3}$, is positive and it is finite (so Z is continuous in σ_y^2). Therefore, Equation 23 implies that $\frac{dZ}{d\sigma_y^2} > 0 \Leftrightarrow (3\tau - \tau\delta - \tau^2\delta - 1) > 0$, i.e., welfare increases with the precision of public information ($\frac{dZ}{d\sigma_y^2} < 0$) if and only if

$$3\tau - 1 < \delta\tau\,(\tau + 1)\,. \tag{24}$$

We now consider three cases. (a) For $0 > \tau > -1$, Inequality 24 holds for $\delta < \frac{3\tau-1}{\tau^2+\tau} \equiv \delta^a(\tau)$. Moreover, with $0 > \tau > -1$, a calculation establishes that $\frac{3\tau-1}{\tau^2+\tau} > 1$; in addition, $\delta < 1$. Therefore, greater precision of public information increases welfare for $0 > \tau > -1$. (b) For $\tau < -1$, Inequality 24 holds for $\delta > \frac{3\tau-1}{\tau^2+\tau}$. This inequality always holds because for $\tau < -1$, $\frac{3\tau-1}{\tau^2+\tau} < 0$; moreover, $\delta > 0$. Therefore, for $\tau < 0$, Inequality 24 is satisfied: an increase in the precision of public information increases welfare. (c) For $\tau > 0$, Inequality 24 holds if and only if $\delta > \frac{3\tau-1}{\tau^2+\tau}$. Because $\delta > 0$, this inequality is always satisfied for $0 < \tau < \frac{1}{3}$. This argument establishes Part (i).

Part (ii). For $\frac{1}{3} < \tau < 1$,

$$\frac{dZ}{d\sigma_y^2} = 0 \iff \delta = \frac{3\tau - 1}{\tau(\tau+1)}$$

so $\delta = \frac{3\tau-1}{\tau(\tau+1)}$ is the unique extreme point of Z with respect to σ_y^2. Moreover, as established in Part (i), Z is increasing in the precision of the public signal for $\delta > \frac{3\tau-1}{\tau(\tau+1)}$ and decreasing in the precision of the public signal for $\delta < \frac{3\tau-1}{\tau(\tau+1)}$. Therefore, Z is minimized with respect to the precision of public information at $\delta = \frac{3\tau-1}{\tau(\tau+1)}$.

Part (iii). Using the definition $r = \frac{\sigma_y}{\sigma_x}$ we can rewrite Equation 22 as

$$Z\,(r;\tau,\sigma_x) = -\frac{1}{2} A{\sigma_x}^2 \frac{r^2}{\left(1 - \tau + r^2\right)^2} \left(\tau^2 - 2\tau + 1 - 2r^2\tau + r^2\right).$$

This equation gives the Bayesian Nash equilibrium payoff as a function of r when we fix τ and σ_x; r increases as we increase σ_y. For fixed σ_x, we see by inspection that $Z\,(0;\tau,\sigma_x^2) = 0$. To evaluate $\lim_{r\to\infty} Z\,(r;\tau,\sigma_x^2)$ we apply L'Hopital's Rule four times (because both the numerator and the denominator are quartics in r) to obtain

$$\lim_{r\to\infty} Z\left(r;\tau,\sigma_x^2\right) = \lim_{r\to\infty} \frac{12r\,(2\tau-1)}{24r} = \tau - \frac{1}{2}.$$

Thus, $\lim_{r\to\infty} Z\left(r;\tau,\sigma_x^2\right) > Z\left(0;\tau,\sigma_x^2\right)$ if and only if $\tau > 0.5$. □

Proof. (Proposition 2) Part i. Using the equilibrium payoff, Equation 22, we have

$$\begin{aligned}\frac{dZ}{d\sigma_x^2} &= -\tfrac{1}{2}A\tfrac{\left(\sigma_y^2\right)^2}{\left(\sigma_x^2+\sigma_y^2-\sigma_x^2\tau\right)^3}\left(\sigma_x^2+\sigma_y^2-\sigma_x^2\tau-2\sigma_y^2\tau\right)\\ &= -\tfrac{1}{2}A\tfrac{(\delta-1)^2}{(1-\tau\delta)^3}\left(\tau\delta-2\tau+1\right).\end{aligned} \tag{25}$$

The coefficient $-\frac{1}{2}A\frac{(\delta-1)^2}{(1-\tau\delta)^3} < 0$. Therefore, welfare increases with the precision of private information (i.e. $\frac{dZ}{d\sigma_x^2} < 0$) if and only if $(\tau\delta - 2\tau + 1) > 0$. This inequality holds for $\tau < 0.5$.

Part ii. From Part (i), the payoff is continuous in σ_x^2, and for $\tau > 0.5$ the unique extreme point of the payoff, with respect to σ_x^2, is $\delta = \frac{2\tau-1}{\tau}$. Moreover, the payoff increases with the precision of the private signal if and only if $\delta > \frac{2\tau-1}{\tau}$, and the payoff decreases with the precision of the private signal if and only if $\delta < \frac{2\tau-1}{\tau}$. Recall that an increase in the precision of the private signal corresponds to an increase in r and a decrease in δ. Thus, for fixed $\tau > 0.5$ and fixed σ_y^2, the payoff is maximized with respect to the precision of private information at $\delta = \frac{2\tau-1}{\tau}$. □

B.2 Analysis with general ρ

This Online Appendix collects the analysis for general values of ρ. Where this analysis requires numerical methods, we use our point estimates $r = 0.923, \rho = 0.247, \tau = 0.448$. The point estimate of ρ is much larger than the theoretical maximum value $\rho = 0.09$ when there are $n = 135$ firms (the median daily number fishing). Recall that the estimation procedure is independent of n; this is a reflection of the fact that the payoff and the equilibrium decision rule are independent of n when there are no clubs ($c = 1$) and no global information sharing ($f = 0$). The introduction to Section 4 explains this point.

B.2.1 Comparative statics

Using Equation 3 with $c = 1$, we obtain

$$\begin{gathered}\text{(i) } \tfrac{d\gamma^{NE,1}}{d\rho} < 0 \Leftrightarrow 1 - r^2 - \tau < 0, \quad \text{(ii) } \tfrac{d\gamma^{NE,1}}{d\tau} < 0 \Leftrightarrow \rho < r, \\ \text{and} \quad \text{(iii) } \tfrac{d\gamma^{NE,1}}{dr} < 0 \Leftrightarrow \left(r^2 - \tau + 1\right)\rho < 2r\left(1 - \tau\right).\end{gathered} \tag{26}$$

In general, the signs of the expressions are ambiguous, but at our point estimates, all three inequalities (after the three "$\Leftrightarrow$"s) are satisfied. Thus, for $c = 1$ and our point estimates, the equilibrium weight on the public signal falls with: greater correlation between the noise in the public and private signal; greater congestion; and a relatively more precise private signal. These results are intuitive. With greater correlation, the public signal adds less new information about θ to the private signal; firms therefore rely less on the public signal in order to increase their distance from other firms, thereby lowering their congestion costs. Greater congestion costs (larger τ) reduces firms' reliance on the public sign for a similar reason. Finally, a relatively more precise private signal shifts firms' equilibrium weight to that signal.

Because $\rho < 1$, the denominator of δ is $1+r^2-2\rho r > 1+r^2-2r = (1-r)^2 \geq 0$. Therefore, δ has the same sign as the numerator, $1 - \rho r$, and δ is continuous in ρ and r. The function is non-monotonic in parameters, with

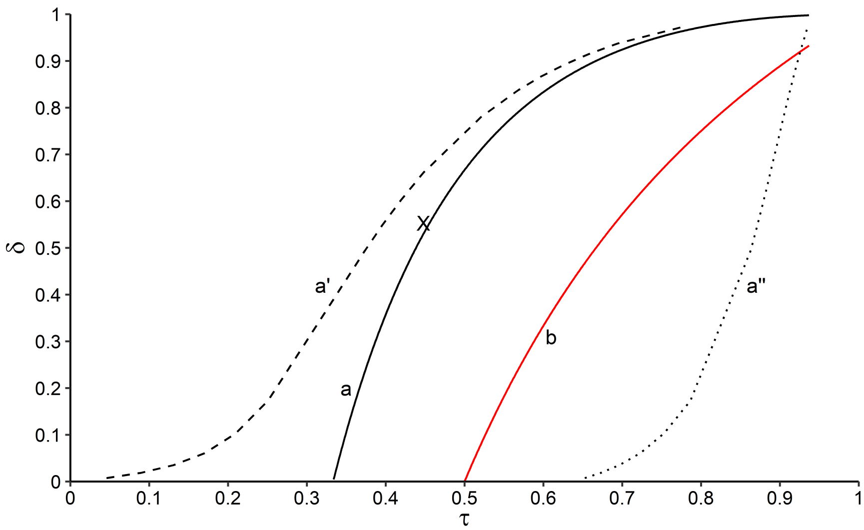
$$\frac{d\delta}{d\rho} < 0 \Leftrightarrow r > 1 \text{ and } \frac{d\delta}{dr} > 0 \Leftrightarrow \rho > 2\frac{r}{r^2+1}. \tag{27}$$

A higher correlation between noise in signals increases the BLUE weight on the public signal if and only if the public signal is relatively precise, compared to the private signal ($r < 1$). For ρ close to 0, a relatively more precise public signal (smaller r) increases the BLUE weight on the public signal, as we would expect. However, for $r > 3$ and sufficiently large ρ, a relatively less precise public signal increases the BLUE weight on the public signal. For $r = 0.923$, we have $\frac{d\delta}{d\rho} > 0$ and $\frac{d\delta}{dr} < 0$.

B.2.2 Constructing the boundaries for welfare effects when $\rho \neq 0$.

For $\rho = 0$, the curves labeled a and b in Figure B1 show the boundaries in the (τ, δ) plane at which the welfare effect of increased precision changes signs. These curves were discussed in Section 4.1 and illustrated in Figure 1. The curves labeled a' and a'' show new boundaries for public information when $\rho = 0.247$. With this value of ρ, a more precise public signal raises welfare for parameters above the curve a' and below the curve a'', and lowers welfare for parameters below those two curves. The larger ρ changes the boundary b for private information

Figure B1: Boundaries that determine the welfare effect of more precise information



Notes: For $\rho = 0,\ c = 1$: (i) more precise public information increases welfare if and only if (τ, δ) lies above the solid curve labelled a; (ii) more precise private information increases welfare if and only if (τ, δ) lies above the solid curve labelled b. For $\rho = 0.247,\ c = 1$: (i) more precise public information increases welfare if and only if (τ, δ) lies above the dashed curve labelled a' or below the dotted curve labelled a''; (ii) the boundary for private information is very close to the curve b (see Figure B2); we do not show that boundary here in order to reduce clutter. Under $\rho = 0.247$, the point X identifies our point estimate $(\tau, \delta) = (0.448, 0.553)$, which lies above the curve a.

only slightly, so to maintain clarity, this figure does not show that new boundary (corresponding to $\rho = 0.247$). We discuss this figure and then explain how we obtain the new boundaries.

The derivatives of welfare with respect to the precision of information are continuous for $\tau < 1 - \rho^2 = 0.94$ (at $\rho = 0.247$). To use the same figure to graph the boundaries with both $\rho = 0$ and $\rho = 0.247$, Figure B1 truncates the boundary curves at $\tau = 0.94$.

The critical τ at which the derivative of welfare with respect to more precise public information vanishes is a solution to a cubic in τ. The curves a' and a'' show the graphs of the two solutions to this cubic for $0 < \tau < 0.94$. The derivative of welfare is negative between these two curves and positive above a' and below a''. In view of the fact that the relevant derivative is a ratios of cubics, this level of complexity is not surprising.

Figure B1 nevertheless conveys a simple message. First, even a substantial correlation between the noise in the public and private signals has almost no effect on the comparative statics of welfare with respect to the precision of private information (i.e., the boundary b scarcely changes). Second, for τ in the range that is plausible for the anchoveta fishery (where likely $\tau < 0.65$), positive correlation between the noise in the signals reduces the parameter space over which more precise public information raises welfare: the curve a' lies above a, and the curve a'' does not exist for $\tau < 0.65$. This ranking could be reversed, but only for unlikely values of τ (in the region below curve a''). Third, taking correlation into account might reverse policy conclusions. The point X in Figure B1, identifying our point estimate of (τ, δ) given $\rho = 0.247$, lies between the curves a' and a. The change in sign of the welfare effect of increased precision of public information compared to when $\rho = 0$ arises for two reasons in this example: (i) we used a large value of ρ, leading to a substantial shift in the boundary a; (ii) our point estimates for (τ, δ), indicated by the X, lies close to the boundary a. On this boundary, the welfare effect of more precise public information is zero, and because of continuity the welfare effect is close to zero near the boundary. Therefore, even a small shift in the boundary can switch the sign of the welfare effect. However, because the magnitude of the welfare effect is small, this flip is unlikely to be economically important.

We now explain how (for $\rho \neq 0$) we use Equations 1, 3 and 4 and numerical methods to construct the boundaries in the (τ, δ) plane where an increase in precision of public or private information has zero welfare effect. We use MuPad, a feature of Scientific Workplace, for calculations. For given ρ, Equation 1 gives the relation between δ and r. When $\rho = 0$, $\delta \in [0, 1]$ as r varies over the positive half line. For $\rho > 0$ we need to restrict the domain of r so that its image $\delta(r; \rho) \in [0, 1]$. We want to keep δ in this range because our goal is to compare (across $\rho = 0$ and $\rho > 0$) the boundaries at which the welfare effect of more precise information changes signs. To this end, for $\rho > 0$, we restrict $r \in R(\rho) \equiv [\rho, \frac{1}{\rho}]$.

We first show that for $\rho > 0$, $\delta(r; \rho)$ is a strictly decreasing function of r for $r \in R(\rho)$. Dropping the argument ρ from the function $\delta(r; \rho)$, note that $\delta(r)_{|r=\rho} = 1$ and $\delta(r)_{r=\frac{1}{\rho}} = 0$. Moreover, using the second inequality in System 27 we confirm that $\frac{d\delta}{dr} < 0$ at the boundaries of the interval $R(\rho)$. Again using System 27, we see that $\frac{d\delta}{dr} = 0$ requires $h(r) \equiv \rho r^2 - 2r + \rho = 0$. The roots of $h(r)$ are $\frac{1 \pm \sqrt{1-\rho^2}}{\rho}$. The smaller root lies below the lower boundary $r = \rho$. To confirm this claim, note that

$$\frac{1 - \sqrt{1-\rho^2}}{\rho} - \rho = \frac{1 - \rho^2 - \sqrt{1-\rho^2}}{\rho} < 0.$$

By inspection, the larger root of $h(r) = 0$ lies above the upper boundary of $R(\rho)$. In summary, we have shown that for $\rho > 0$, $\delta(r)$ is a strictly decreasing function of r for $r \in R(\rho)$.[24]

For completeness, we note that for $\rho < 0$, $\delta(r)$ is strictly decreasing for $r \in [0, \infty]$, with $\delta(r)_{|r=0} = 1$ and $lim_{r \to \infty} \delta(r) = 0$.

[24]Denote $r^- \equiv \frac{1-\sqrt{1-\rho^2}}{\rho}$, the smaller root of $h(r)$. For $\rho > 0$ and $r \in [0, \rho)$, $\delta(r)$ is non-monotonic, reaching its maximum at $r = r^-$. For $r > \frac{1}{\rho}$, $\delta(r)$ is non-monotonic, reaching its minimum at $r^+ \equiv \frac{1+\sqrt{1-\rho^2}}{\rho}$, the larger root of $h(r)$.

To ensure that our domain of analysis remains $0 < \delta < 1$ when $\rho \neq 0$ we adopt

Assumption 2. *For $\rho > 0$, $r \in R(\rho) \equiv \left(\rho, \frac{1}{\rho}\right)$, and for $\rho \leq 0$, $r \in (0, \infty)$.*

We find the boundaries in the (τ, δ) plane where the derivative of the payoff with respect to the precision of either the public or the private signal equals zero in two steps. In the first step we find the boundaries in the (τ, r) plane where the derivative of the payoff with respect to the precision of either the public or the private signal equals zero. We then convert these boundaries from the (τ, r) plane to the (τ, δ) plane, our ultimate objective.

To avoid having to introduce yet another symbol, we abuse notation by using S (only in this Online Appendix) to mean the function S defined in Equation 4 after replacing γ with the equilibrium weight given in Equation 3. With this understanding, we have

$$S = r^2 \frac{2\tau + r^2\rho^2 + 2r\rho - \tau^2 + \rho^2 + 2r^2\tau - 2r\rho^3 - r^2 - 4r\tau\rho - 1}{\left(-\tau - 2r\rho + r^2 + 1\right)^2}.$$

Using MuPad, we can write the derivative as

$$\frac{dS}{dr} = \frac{N}{p^3}, \tag{28}$$

with the following definitions

$$\begin{aligned} N &\equiv c_0 + c_1\rho + c_2\rho^2 + c_3\rho^3 + c_4\rho^4 \\ p &\equiv -r^2 + 2\rho r + \tau - 1, \end{aligned} \tag{29}$$

together with

$$\begin{aligned} c_0 &\equiv 2r\left(3r^2\tau^2 - 4r^2\tau + r^2 - \tau^3 + 3\tau^2 - 3\tau + 1\right) \\ c_1 &\equiv \left(-2r\left(r^3 - 2r^3\tau + 6r\tau^2 - 9r\tau + 3r\right)\right) \\ c_2 &\equiv 2r\left(\tau - 2r^2\tau + r^2 - 1\right) \\ c_3 &\equiv 2r\left(3r - 3r\tau + r^3\right) \\ c_4 &\equiv \left(-4r^3\right). \end{aligned} \tag{30}$$

In this Online Appendix (only) we use N as a mnemonic for "numerator" and we use p as a mnemonic for "pole", because the derivative is undefined where $p = 0$, i.e., at a pole.

The derivative of the payoff with respect to the precision of public information has the same sign as $-\frac{dS}{dr} = \frac{N}{-p^3}$. The derivative changes discontinuously, flipping signs as τ passes through a pole, defined as a value of τ where $p = 0$. The unique solution to $p = 0$ (the only pole) is $\tau^p \equiv r^2 - 2\rho r + 1$. The minimum of this function occurs at $r = \rho$, where $\tau = 1 - \rho^2$.

Our model already requires $\tau < 1$. We further restrict the domain of analysis to $\tau < 1 - \rho^2$, thus ensuring that the domain includes no poles. With this restriction, the derivative of the payoff with respect to the precision of public information is continuous. For our preferred point estimate $\rho = 0.247$, this restriction eliminates only $\tau \in (0.938\,99, 1)$ This reduction in the domain is unimportant in our setting, where our point estimates imply much smaller values of τ.

Because $\rho < 1$, we have

$$p_{|\tau=0} = -\left(1 - 2\rho r + r^2\right) < -\left(1 - 2r + r^2\right) = -\left(1 - r\right)^2 \leq 0,$$

so $p_{|\tau=0} < 0$; moreover, p is strictly increasing in τ, and equals 0 at $\tau = \tau^p$. Our restriction to $\tau < 1 - \rho^2$ implies that $p < 0$ for $\tau < 1 - \rho^2$, so $-p^3$, the demoninator of $-\frac{dS}{dr}$, is positive over our domain of analysis. The derivative of the payoff with respect to the precision of the public signal therefore has the same sign as n over this domain. By inspection of the definitions in System 29, n is a cubic in τ.

Now we establish that $N_{|\tau=0} > 0$. We have

$$N_{|\tau=0} = 2r\left(1 - \rho^2\right)\left(-r^3\rho + 2r^2\rho^2 + r^2 - 3r\rho + 1\right). \tag{31}$$

The coefficient $2r\left(1 - \rho^2\right) > 0$, so $n_{|\tau=0}$ has the same sign as

$$f \equiv -r^3\rho + 2r^2\rho^2 + r^2 - 3r\rho + 1.$$

We now establish that for $r \in R(\rho) = \left(\rho, \frac{1}{\rho}\right)$ when $\rho > 0$ or $r \in [0, \infty]$ when $\rho \leq 0$, $f > 0$. This inequality and Equation 31 implies that $N_{|\tau=0} > 0$. Note that $\frac{df}{d\rho} = -r\left(r^2 - 4\rho r + 3\right)$. Also, the roots of $f = 0$ are $\left\{\frac{1}{r}, \frac{1}{2r}\left(r^2 + 1\right)\right\}$ for $r \neq 0$. Both of these roots are positive. Using $f_{|\rho=0} = r^2 + 1 > 0$ and $\frac{df}{d\rho}_{|\rho=0} = -r\left(r^2 + 3\right) < 0$, and the fact that both roots of $f = 0$ are positive, we conclude that $f > 0$ for $\rho \leq 0$.

Now consider the case $\rho > 0$. We have $f_{|\rho=0} = r^2 + 1 > 0$ and $f_{|\rho=r} = \left(r^2 - 1\right)^2 > 0$. We need to examine two possibilities: $\frac{1}{r} < \frac{1}{2r}\left(r^2 + 1\right)$ and $\frac{1}{r} > \frac{1}{2r}\left(r^2 + 1\right)$. The first of these two case implies $2 < r^2 + 1$, or $1 < r^2$. We fix r and consider variations in ρ. We have $f(\rho; r) > 0$ for $\rho < \frac{1}{r}$, i.e., for $r < \frac{1}{\rho}$. This inequality holds because of our assumption that $r \in R(\rho)$. Now consider the second case, where $\frac{1}{r} > \frac{1}{2r}\left(r^2 + 1\right)$, i.e., $2 > r^2 + 1$, or $1 > r^2$. This inequality implies $\frac{r^2+1}{2r} > r$. Therefore, there exists no $\rho > 0$ where $\rho \in \left(\frac{r^2+1}{2r}, r\right)$, i.e. there exists no $\rho < r$ (as required by $r \in R(\rho)$) where $\rho > \frac{r^2+1}{2r}$ (a necessary condition for $f < 0$).

We conclude that under Assumption 2, $N_{|\tau=0} > 0$, so the derivative of the payoff with respect to the precision of public information is positive at $\tau = 0$. In addition, the derivative has the same sign as the cubic $N(\rho, r, \tau)$ for $\tau < 1-\rho^2$. These two facts enable us to determine the boundaries at which the derivative changes signs by finding the roots of $N(\rho, r, \tau) = 0$; we are interested only in real roots less than $\tau = 1 - \rho^2$.

Before turning to the effect of an increase in the precision of private information, we note as a consistency check that

$$-\frac{dS}{dr}_{|\rho=0} = -2r\frac{\tau-1}{(r^2-\tau+1)^3}\left(-3r^2\tau + r^2 + \tau^2 - 2\tau + 1\right).$$

The term $-2r\frac{\tau-1}{(r^2-\tau+1)^3} > 0$ so the derivative has the same sign as

$$\left(-3r^2\tau + r^2 + \tau^2 - 2\tau + 1\right).$$

Comparing this expression to the first line of Equation 23 we see that our general expression (that is, for $\rho \in (-1, 1)$) for the boundaries where the derivative of the payoff with respect to the precision of public information changes signs specializes to the expression that we obtained in Proposition 1 where $\rho = 0$.

To determine *the derivative of the payoff with respect to the precision of private information*, we begin with the payoff $P = \frac{A}{2}\sigma_x^2 S$. The welfare effect of an increase in precision of the private information (a reduction in σ_x^2) is

$$\begin{aligned} -\tfrac{dP}{d\sigma_x} &= -\left(A\sigma_x S - \tfrac{A}{2}\sigma_x^2 \tfrac{dS}{dr}\tfrac{\sigma_y}{\sigma_x^2}\right) = -\tfrac{A}{2}\sigma_x\left(2S - \sigma_x \tfrac{dS}{dr}\tfrac{\sigma_y}{\sigma_x^2}\right) \\ &= -\tfrac{A}{2}\sigma_x\left(2S - \tfrac{dS}{dr}r\right) = -\tfrac{A}{2}\sigma_x\left(2S - \tfrac{n}{p^3}r\right). \end{aligned}$$

The first equality uses the definition of r and the last equality uses Equation 28. Using MuPad we obtain

$$-\left(2S - \frac{n}{p^3}r\right) = 2r^3\frac{r-\rho}{-p^3}m$$

with

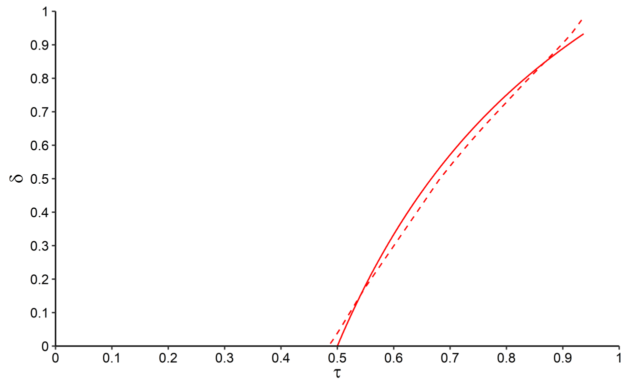
$$m \equiv \left(4r\rho - 2r^2 - \rho^2 - 1\right)\tau + \left(1-\rho^2\right)r^2 + \left(\rho^2 - 1\right)2\rho r + \left(1-\rho^2\right).$$

Because $2r^3\frac{r-\rho}{-p^3} > 0$ over our domain of analysis, we conclude that the sign of the derivative of welfare with respect to the precision of private information is the same as the sign of m. We have

$$m_{|\tau=0} = \left(1-\rho^2\right)r^2 + \left(\rho^2-1\right)2\rho r + \left(1-\rho^2\right).$$

This expression is positive at $r = 0$ and there are no real roots of $(1-\rho^2)\,r^2 + (\rho^2-1)\,2\rho r + (1-\rho^2) = 0$, so we conclude that $m_{|\tau=0} > 0$. The coefficient of τ, $(4r\rho - 2r^2 - \rho^2 - 1)$ is negative at $r = \rho$ and the derivative of this coefficient with respect to r is $4\rho - 4r < 0$ for

Figure B2: Graphs of the boundaries for $\rho = 0.247$ and $\rho = 0$



Notes: For $\rho = 0.247$ and $c = 1$, an increase in the precision of private information raises welfare if and only if (τ, δ) lies to the left of the dashed boundary. The solid curve shows the boundary that corresponds to $\rho = 0$ and $c = 1$, obtained from Proposition 2 and displayed in Figures 1 and B1. Because the two curves lie so close together, we do not include the dashed graph of the boundary when $\rho = 0.247$ in Figure B1.

$r \in R(\rho)$; therefore, the coefficient is negative over the region of analysis. The unique root of $m = 0$ is

$$b \equiv -\frac{\left(\rho^2 + r^2\left(\rho^2 - 1\right) - 2r\rho\left(\rho^2 - 1\right) - 1\right)}{-4r\rho + \rho^2 + 2r^2 + 1}. \tag{32}$$

From the previous analysis, the derivative of the payoff with respect to the precision of private information is positive for $\tau < b$ and negative for $b < \tau < 1 - \rho^2$.

As a consistency check, we note that when we set $\rho = 0$, b specializes to the boundary given in Proposition 2 (which assumes $\rho = 0$). Specializing the formula for b by setting $\rho = 0$ we have $b_{|\rho=0} = \frac{r^2+1}{2r^2+1}$. In Proposition 2 we expressed the boundary as $\delta = \frac{2\tau-1}{\tau}$. Using $\delta_{|\rho=0} = \frac{1}{r^2+1}$ we rewrite this boundary as $\frac{1}{r^2+1} = \frac{2\tau-1}{\tau}$. Solving for τ gives $\tau = \frac{r^2+1}{2r^2+1} = b_{|\rho=0}$, thus confirming that our formula for general ρ specializes to the formula in Proposition 2 when $\rho = 0$.

Algorithm for finding the boundaries in Figures B1 and B2. We have shown that the boundaries for public information are real roots of the cubic $n(\tau; \rho, r) = 0$ that satisfy $\tau < 1 - \rho^2$; the unique boundary for private information is the function $b(r, \rho)$. For fixed ρ, both of these boundaries are expressed as functions of r. To convert these to functions of δ (as we require for the two figures) we use the relation $\delta = \frac{1-r\rho}{r^2-2\rho r+1}$. For $\rho > 0$ we proceed as follows:

(i) Choose a value ρ.

(ii) Select a grid over $R(\rho) = \left(\rho, \frac{1}{\rho}\right)$ by choosing a set of increasing weights w that satisfy $0 < w < 1$, and for each weight set $r(w, \rho) = \rho + w\left(\frac{1}{\rho} - \rho\right)$.

(iii) For this combination (ρ, w) find the real roots of $n(\tau; \rho, r(w, \rho)) = 0$ that satisfy

Table B1: Boundary values for public and private information

(i) w	(ii) r	(iii) δ	(iv) τ^l	(v) τ^u	(vi) b
0.02	0.3230	0.9740	0.7799	0.9355	0.9333
0.04	0.3991	0.9369	0.6938	0.9287	0.9170
0.05	0.4371	0.9148	0.6576	0.9250	0.9054
0.06	0.4751	0.8907	0.6249	0.9210	0.8922
0.07	0.5131	0.8648	0.5951	0.9170	0.8775
0.08	0.5511	0.8375	0.5680	0.9129	0.8617
0.10	0.6272	0.7800	0.5208	0.9046	0.8285
0.14	0.7792	0.6607	0.4478	0.8879	0.7623
0.16	0.8553	0.6026	0.4194	0.8795	0.7321
0.18	0.9313	0.5472	0.3952	0.8712	0.7046
0.20	1.0073	0.4952	0.3742	0.8630	0.6799
0.40	1.7676	0.1733	0.2509	0.7866	0.5487
0.50	2.1478	0.1031	0.2099	0.7531	0.5235
0.60	2.5279	0.0612	0.1709	0.7230	0.5084
0.70	2.9081	0.0351	0.1311	0.6963	0.4987
0.80	3.2883	0.0184	0.0896	0.6729	0.4922
0.90	3.6684	0.0074	0.0459	0.6524	0.4876

Notes: Column (i) contains the weights, w, used to determine the grid points $r \in R(\rho)$. Column (ii) contains the corresponding value of r, given $\rho = 0.247$. Column (iii) contains the corresponding δ. Columns (iv) and (v) contain the lower and upper boundaries, denoted τ^l and τ^u, at which the derivative of welfare with respect to the precision of public information equals 0. More precise public information increases welfare for $\tau < \tau^l$ and for $\tau^u < \tau < 1-\rho^2$, and lowers welfare for $\tau^l < \tau < \tau^u$. Column (vi) contains b, the boundary above which more precise private information raises welfare.

$\tau < 1 - \rho^2$. These are the boundaries for public information. Find $b\left(r\left(w,\rho\right),\rho\right)$; this is the boundary for private information.

Table B1 shows the results of applying this algorithm, using $\rho = 0.247$ and a set of weights w. For example, for $w = 0.16$ (in column (i)) we obtain $r\left(w,\rho\right) = \rho + w\left(\frac{1}{\rho} - \rho\right) = 0.8553$ (in column (ii)), and $\delta = 0.6026$ (in column (iii)). The roots of $n = 0$ are (approximately) $\{0.4194, 0.8795, 2.628\}$. The smallest root is $\tau^l = 0.4194$ (in column (iv)) and the middle root is $\tau^u = 0.8795$ (in column (v)); we ignore the largest root, which lies above the upper bound of our domain of analysis, $1 - \rho^2$). The boundary for private information is $b = 0.7321$. The curves a' and a'' in Figure B1 correspond to columns (iv) and (v) in Table B1, and the dashed curve in Figure B2 corresponds to column (vi) in this table.

B.2.3 Computation of elasticities α and β

We now derive the elasticies α and β. Using the definition of S in Equation 4, we obtain (with the assistance of Mupad, a feature of Scientific Workplace) the closed form expression for the elasticity of S with respect to r:

$$\frac{dS}{dr}\frac{r}{S} = -2\frac{b_0 + b_1\rho + b_2\rho^2 + b_3\rho^3 + b_4\rho^4}{a_0 + a_1\rho + +a_2\rho^2 + a_3\rho^3 + a_4\rho^4}, \tag{33}$$

with

$$b_0 = \left(3r^2\tau^2 - 4r^2\tau + r^2 - \tau^3 + 3\tau^2 - 3\tau + 1\right)$$
$$b_1 = \left(2r^3\tau - r^3 - 6r\tau^2 + 9r\tau - 3r\right)$$
$$b_2 = \left(\tau - 2r^2\tau + r^2 - 1\right)$$
$$b_3 = \left(3r - 3r\tau + r^3\right)$$
$$b_4 = \left(-2r^2\right)$$

and

$$a_0 = -\left(r^2 - \tau + 1\right)\left(r^2 - 2r^2\tau + \tau^2 - 2\tau + 1\right)$$
$$a_1 = \left(2r\left(r^2 - 2r^2\tau + \tau^2 - 2\tau + 1\right) + \left(2r - 4r\tau\right)\left(r^2 - \tau + 1\right)\right)$$
$$a_2 = \left(\left(r^2 + 1\right)\left(r^2 - \tau + 1\right) - 2r\left(2r - 4r\tau\right)\right)$$
$$a_3 = \left(-2r\left(r^2 - \tau + 1\right) - 2r\left(r^2 + 1\right)\right)$$
$$a_4 = \left(4r^2\right).$$

The sign of S is easy to evaluate, given r, τ, ρ. With this information and using the definitions in the first line of Equation 6 we obtain α, the elasticity of the payoff with respect to the precision of public information. We set $\rho = 0$ to obtain the second line of Equation 6:

$$\frac{dS}{dr}\frac{r}{S} = -2\frac{b_0}{a_0}.$$

To obtain β, the elasticity of the payoff with respect to the precision of private information, we use the fact that σ_x affects the payoff, $P = \frac{A}{2}\sigma_x^2 S$, via $S(r(\sigma_x))$ and also via the coefficient σ_x^2. To evaluate the elasticity, we use the formula for the elasticity of σ^2 with respect to $\frac{1}{\sigma}$:

$$\frac{d\sigma^2}{d\frac{1}{\sigma}}\frac{\frac{1}{\sigma}}{\sigma^2} = \frac{d\sigma^2}{d\sigma}\frac{d\sigma}{d\frac{1}{\sigma}}\sigma^{-3} = 2\sigma\sigma^{-3}\left[\frac{d\frac{1}{\sigma}}{d\sigma}\right]^{-1} = -2\sigma\sigma^{-3}\left[\sigma^{-2}\right]^{-1} = -2. \tag{34}$$

Using the definition $r = \frac{\sigma_y}{\sigma_x}$ we have the elasticity

$$\frac{dr}{d\frac{1}{\sigma_x}}\frac{\frac{1}{\sigma_x}}{r} = 1. \tag{35}$$

We use these intermediate results and Equation 33 (the elasticity of S with respect to r) to obtain the elasticity of the payoff with respect to $\frac{1}{\sigma_x}$, β. We have

$$\begin{aligned}\beta \equiv \frac{dP}{d\left(\frac{1}{\sigma_x}\right)}\frac{\frac{1}{\sigma_x}}{|P|} &= \frac{A}{2}S\frac{d\sigma_x^2}{d\frac{1}{\sigma_x}}\left[\frac{\frac{1}{\sigma_x}}{\sigma_x^2}\right]\frac{\sigma_x^2}{\frac{A}{2}\sigma_x^2|S|} + \frac{A}{2}\sigma_x^2\frac{dS}{dr}\frac{r}{|S|}\frac{|S|}{r}\frac{dr}{d\frac{1}{\sigma_x}}\frac{\frac{1}{\sigma_x}}{\frac{A}{2}\sigma_x^2|S|} \\ &= -2\frac{S}{|S|} - \alpha.\end{aligned} \tag{36}$$

The first term after the first equality gives the percentage change in the payoff (Equation 4) due to a change in the coefficient (σ_x^2) of S. The second term gives the percentage payoff change operating through the change in r. The second line follows from simplification, using Equation 34 for the first term and Equation 35 for the second term. In particular, the second term simplifies to $\frac{dS}{dr}\frac{dr}{d|S|}$, which equals $-\alpha$ by Equation 6. To obtain Equation 7, we transform $-2\frac{S}{|S|}$ into $sign(-P)\cdot 2$ using $\frac{S}{|S|} = -sign(-P)1$. ($P < 0 \iff S < 0$.)

B.3 The comparative statics of clubs

Using Equation 3 we obtain the derivative of the Bayesian Nash equilibrium weight with respect to the club size

$$\frac{d\gamma^{NE,c}}{dc} = r\left(r-\rho\right)\left(n-1\right)\frac{n\tau - n + 1}{\left(n + c\tau - n\tau - cr^2 + 2cr\rho + cnr^2 - 2cnr\rho - 1\right)^2}.$$

The right side of this equality has the same sign as $(r-\rho)\left(1 - n\left(1-\tau\right)\right)$.

To produce Figure 4, we use the expression for the equilibrium payoff and decision rule from Lemma 1. Dividing this expression by $\frac{A}{2}\sigma_x^2$, we obtain the normalized payoff, S. We evaluate this payoff at our point estimates, $(\tau, r) = (0.448, 0.923)$ and use $c = 1 + (1-\phi)n$ to eliminate c. The result is a ratio of polynomials, a function of (ϕ, n) that we denote T:

$$T \equiv -1.0\frac{(81.288n - 59.915)\phi^2 + (9.4352n + 124.52)\phi + 41.883}{((90.723n2 - 86.03n + 20.395)\phi^2 + (299.01n - 141.77)\phi + 246.38} \tag{37}$$

Figure 4 graphs this function over ϕ for three values of n.

B.4 Global information sharing

To consider the case where firms share some, but not necessarily all of their private information, we assume that the representative firm, i, gets the public signal, $y = \theta + \varepsilon_y$, and m private signals, $(x_{i,1}, x_{i,2}, ...x_{i,m})$, with $x_{i,h} = \theta + \varepsilon_{i,h}$. The expectation of all errors is zero (so the signals are unbiased), and $\mathbb{E}\varepsilon_y^2 = \sigma_y^2$, $\mathbb{E}\left(\varepsilon_{i,h}\right)^2 = \tilde{\sigma}_x^2$, $\mathbb{E}\varepsilon_y\varepsilon_{i,h} = \psi\tilde{\sigma}_x\sigma_y$, $\mathbb{E}\varepsilon_{i,h}\varepsilon_{j,n} = 0$ for all i, j, h, n. The correlation between noise in any of the firm's private signals and the noise in the public signal is ψ. The noise is uncorrelated across i's signals and also uncorrelated with noise in every other firm's private signals. (We already assumed that the noise in private signals is uncorrelated across firms; the assumption that noise in private signals are uncorrelated within a firm is new.) Because all of i's signals are equally informative, the firm puts equal weight on each. Define the average of i's private signals as $\bar{x}_i \equiv \frac{1}{m}\sum_{k=1}^{m} x_{i,k}$. The above assumptions imply that $\mathbb{E}\left(\bar{x}_i - \theta\right)^2 = \frac{1}{m}\tilde{\sigma}_x^2$ and

$$\mathbb{E}\left(\bar{x}_i - \theta\right)\left(y - \theta\right) = \frac{1}{m}\mathbb{E}\left[\left(\sum_{k=1}^{m} \varepsilon_{i,k}\right)\varepsilon_y\right] = \frac{1}{m} m\psi\tilde{\sigma}_x\sigma_y = \psi\tilde{\sigma}_x\sigma_y.$$

Our baseline model, in which each firm receives a single private signal, denoted x_i, assumed $\mathbb{E}\left(x_i - \theta\right)^2 = \sigma_x^2$ and $\mathbb{E}\left(x_i - \theta\right)\left(y - \theta\right) = \rho\sigma_x\sigma_y$. To make the assumptions across the two models mutually consistent, we set

$$\frac{1}{m}\tilde{\sigma}_x^2 = \sigma_x^2, \text{ or } \tilde{\sigma}_x^2 = m\sigma_x^2 \implies \tilde{\sigma}_x = \sqrt{m}\sigma_x. \tag{38}$$

We also have

$$\psi\tilde{\sigma}_x\sigma_y = \rho\sigma_x\sigma_y, \text{ or } \psi\sqrt{m}\sigma_x\sigma_y = \rho\sigma_x\sigma_y \implies \psi\sqrt{m} = \rho. \tag{39}$$

Now suppose that the regulator requires each firm to reveal $k \leq m$ pieces of private information. With n firms each revealing k unbiased signals whose noise is mutually uncorrelated, the mean of firms' revealed (previously private) information is $\bar{x} \equiv \frac{1}{nk}\sum_{i=1}^{n}\sum_{p=1}^{k} x_{i,p}$. The variance of $\bar{x}$ is

$$\mathbb{E}\left(\frac{1}{nk}\sum_{i=1}^{n}\sum_{p=1}^{k} x_{i,p} - \theta\right)^2 = \frac{1}{nk}\tilde{\sigma}_x^2 = \frac{m}{nk}\sigma_x^2 = \frac{\sigma_x^2}{nf},$$

where we use $f = \frac{k}{m}$, the fraction of its private information that each firm reveals to the regulator. The covariance between the (old) public signal and the aggregation of the regulator's new information is

$$\begin{aligned} cov(\bar{x}, y) &= \mathbb{E}\left(\bar{x} - \theta\right)\left(y - \theta\right) = \mathbb{E}\tfrac{1}{nk}\textstyle\sum_{i=1}^{n}\sum_{p=1}^{k}\left(\varepsilon_{i,p}\varepsilon_y\right) \\ &= \psi\tilde{\sigma}_x\sigma_y = \psi\sqrt{m}\sigma_x\sigma_y = \rho\sigma_x\sigma_y. \end{aligned}$$

The covariance between the regulator's new signal (consisting of the firms' erstwhile private information), $\bar{x}$, and the original public signal, y, is the same as the covariance between the firm's (original) full private signal, x_i, and y. The correlation between $\bar{x}$ and y is

$$corr\left(\bar{x}, y\right) = \frac{\rho\sigma_x\sigma_y}{\sqrt{\frac{1}{nf}}\sigma_x\sigma_y} = \rho\sqrt{nf}.$$

After the partial revelation of private information, the distribution of the regulator's information state, $(y, \bar{x})$, is

$$(y, \bar{x}) \sim \left((\theta, \theta), \begin{bmatrix} \sigma_y^2 & \rho\sigma_y\sigma_x \\ \rho\sigma_y\sigma_x & \frac{\sigma_x^2}{nf} \end{bmatrix} \right).$$

This covariance matrix is positive definite because

$$|\rho| < \frac{1}{\sqrt{n}} < \frac{1}{\sqrt{nf}}. \tag{40}$$

The regulator's BLU estimator of θ, $\lambda y + (1-\lambda)\bar{x}$ (also, its Bayesian posterior given normality and diffuse priors), minimizes the variance of the estimator. It solves

$$\begin{gathered} V(n) = \min_\lambda \left(\lambda^2\sigma_y^2 + (1-\lambda)^2 \frac{\sigma_x^2}{nf} + 2(1-\lambda)\lambda\rho\sigma_y\sigma_x \right) = \\ \min_\lambda \left(\lambda^2 r^2 + (1-\lambda)^2 \frac{1}{nf} + 2(1-\lambda)\lambda\rho r \right)\sigma_x^2. \end{gathered} \tag{41}$$

The solution to this problem gives

$$\lambda = \frac{1 - fnr\rho}{fnr^2 - 2fnr\rho + 1} \text{ and } V(n) = r^2 \frac{1 - fn\rho^2}{fnr^2 - 2fn\rho r + 1}\sigma_x^2. \tag{42}$$

As a consistency check, note that Equation 40 implies that the numerator of $V(n)$ is positive. For $\rho \leq 0$ the denominator is positive by inspection. We therefore consider the case $\rho > 0$, where the denominator is larger than $h(r) \equiv fnr^2 - 2f\sqrt{n}r + 1$. By inspection, $h(0) > 0$, and using the quadratic formula we see that for $0 < f < 1$ there are no real roots to $h(r) = 0$; thus, $h(r) > 0$ for $0 < f < 1$. By inspection, the denominator is positive for $f = 0$. Finally, for $f = 1$ (and $\rho > 0$) the denominator is greater than $k(r) \equiv nr^2 - 2\sqrt{n}r + 1$. We see that $k(0) = 1$ and the unique root of $k(r) = 0$ is $\frac{1}{\sqrt{n}}$, so $k(r) \geq 0$.

Before having received the firms' private information, the variance of the public signal was $\sigma_y^2 = r^2\sigma_x^2$. Therefore, the fractional reduction in the variance due to the additional (formerly private) information is

$$\varpi \equiv \frac{r^2\sigma_x^2 - r^2\sigma_x^2 \frac{1-fn\rho^2}{fnr^2 - 2fn\rho r + 1}}{r^2\sigma_x^2} = \frac{(r-\rho)^2}{r^2 - 2\rho r + \frac{1}{fn}}, \tag{43}$$

shown in the text (when $\rho = 0$) as Equation 8.

We now determine the relation between the post-information-transfer moments and the original moments. The variance of the original public signal is σ_y^2. Using Equation 43, the transfer of information reduces this variance by the fraction ϖ, so the post-transfer variance of the (new) public signal is $(1-\varpi)\sigma_y^2$. Each firm's original private signal consisted of

m independent signals, each with variance $\tilde{\sigma}_x^2 = m\sigma_x^2$. After communicating k pieces of information, the firm has $m-k$ pieces of information that remain private. Because these individual signals are independent, their mean is a sufficient statistic for the firm's (newly reduced) private information. This statistic is

$$\tilde{x}_i \equiv \frac{1}{m-k}\sum_{s=1}^{m-k} x_{is} \sim \left(\theta, \frac{1}{m-k}\tilde{\sigma}_x^2\right) = \left(\theta, \frac{m}{m-k}\sigma_x^2\right).$$

where the equality uses $\tilde{\sigma}_x^2 = m\sigma_x^2$ (Equation 38). Using $\frac{k}{m} = f$, we write the variance of this signal as $\frac{1}{1-f}\sigma_x^2$. After the information transfer, the ratio of the standard deviations of the (new) private and public information is

$$r' = \frac{\sqrt{1-\varpi}\sigma_y}{\sqrt{\frac{1}{1-f}}\sigma_x} = r\sqrt{(1-\varpi)(1-f)}.$$

Equation 9 collects the expressions for the variances of the "new" public and private signals, and the ratio of their standard deviations.

We need an intermediate result to express the covariance between the new (after the transfer of information) public and private signals in terms of ϖ. Evaluating $\lambda = \frac{1-fnr\rho}{fnr^2-2fnr\rho+1}$ (from Equation 42) at $n = \frac{\varpi}{\left((r-\rho)^2-\varpi(r^2-2r\rho)\right)f}$ gives

$$\lambda = \frac{1}{r-\rho}\left(r\left(1-\varpi\right)-\rho\right). \tag{44}$$

The covariance between $\tilde{x}_i$ and the (new) public signal is

$$\begin{gathered}\mathbb{E}\left(\frac{1}{m-k}\sum_{s=1}^{m-k}\varepsilon_{i,s}\right)\left(\lambda\varepsilon_y + (1-\lambda)\frac{1}{nk}\sum_{i=1}^{n}\sum_{p=1}^{k}\varepsilon_{i,p}\right) = \\ \mathbb{E}\left(\frac{1}{m-k}\sum_{s=1}^{m-k}\varepsilon_{i,s}\right)\lambda\varepsilon_y = \lambda\rho\sigma_x\sigma_y = \frac{\rho}{r-\rho}\left(r\left(1-\varpi\right)-\rho\right)\sigma_x\sigma_y.\end{gathered} \tag{45}$$

The left side of the first equation uses the formula for the revised public signal, after the regulator aggregates information from firms. The first equality uses the assumption that noise in the firm's remaining private information is uncorrelated with noise in all other (previously) private pieces of information that were communicated to the regulator. The second equality uses the fact (established above) that our assumptions imply that $\mathbb{E}\varepsilon_{i,s}\varepsilon_y = \rho\sigma_x\sigma_y$. The final equality uses Equation 44. The correlation between the new public and private signals is

$$\rho' = \frac{\frac{1}{r-\rho}\left(r\left(1-\varpi\right)-\rho\right)\rho\sigma_y\sigma_x}{\sqrt{1-\varpi}\sigma_y\sqrt{\frac{1}{1-f}}\sigma_x} = \frac{\frac{1}{r-\rho}\left(r\left(1-\varpi\right)-\rho\right)\rho}{\sqrt{1-\varpi}\sqrt{\frac{1}{1-f}}}. \tag{46}$$

C Online Appendix: Data and empirical analysis

Online Appendix C.1 lists the sets we drop from the electronic logbook data and details the provenance and content of our public information satellite data. Online Appendices C.2 and C.3 specify the construction of the public and private signal matrices that we use to estimate the parameter r. Online Appendix C.4 details the cross-validation procedure involved in estimating r. Online Appendix C.5 assesses the robustness of our estimates of r to alternative specifications. Online Appendix C.6 does the same for our estimation of τ.

C.1 Electronic logbook data cleaning and public information satellite data

The raw electronic logbook data contains 247,024 sets. We drop 97 sets that occur on fishing trips that last longer than two weeks (because the median trip lasts 17 hours and the 99th percentile trip lasts 3 days), as well as 7 sets where tons caught is more than 10 times the 99th percentile of tons caught per set by that vessel. The electronic logbook data we use in our analysis therefore contains 246,920 sets.

We use chlorophyll, daytime sea surface temperature (SST), and nighttime SST from the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument aboard the Suomi National Polar-orbiting Partnership (SNPP) satellite (NASA, 2016a, 2016b, 2018); chlorophyll, daytime SST, and nighttime SST from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor aboard the Aqua satellite (NASA, 2014, 2019a, 2019b); SST anomaly from the National Oceanic and Atmospheric Administration (NOAA) via Instituto Humboldt, a Peruvian non-governmental organization (NOAA, 2022); sea surface salinity from the Soil Moisture Active Passive (SMAP) satellite (NOAA, 2021); and sea level anomaly from the NOAA Laboratory for Satellite Altimetry (NOAA, 2023). We download the data occurring inside Peru's Exclusive Economic Zone (Flanders Marine Institute, 2019). We use the previous day's value for both nighttime SST variables since a firm deciding where to fish today only knows last night's SST value. The resolution of chlorophyll, daytime SST, and nighttime SST data is 4 km (for both VIIRS and MODIS); the resolution of sea surface salinity and SST anomaly is 25 km; and the resolution of sea level anomaly is 0.25° (about 30 km).

C.2 Construction of public signal matrix

First, for each of the nine oceanic geophysical variables, we record the value of the geophysical variable at the nearest location to set i. (Recall that most satellites image only part of the earth each day.) Second, we interpolate each geophysical variable over the entire fishery each day using inverse distance weighting. Then we extract the interpolated value of the geophysical variable for each set. We thus create two variables from each geophysical variable – the nearest value and the interpolated value – which gives us 18 geophysical variables (9 times 2).

Then we calculate three new variables for each of these 18 original variables. We calculate the average value for set i's best local zone, the difference between set i's value and the average value in its best local zone, and the difference between set i's value and the average value that day (in any zone). After this procedure we have 72 predictor variables (18 plus 3 times 18).

Next, we add 18 more predictor variables to reach a total of 90. First, we record the distance between set i and the nearest location at which each of the 9 original, non-interpolated geophysical variables were measured, since closer measurements may be more predictive. Second, we calculate the average of these distances in set i's best local zone.

We then calculate the square of each of these 90 predictor variables, increasing the number to 180. (Including additional higher order terms causes our estimation procedure to exceed the RAM of our server.) Finally, we interact all 180 of these variables with each other, obtaining a public signal matrix with 16,290 variables (180 choose 2, plus the original 180 variables). If a geophysical variable has no measurements inside Peru's anchoveta fishery on a given day, sets on this day have missing values for all variables that are a function of that geophysical variable.

C.3 Construction of private signal matrix

First, for each set i by a vessel in firm f on day t, we record the CPUE of the nearest set on day $t-1$ by vessels in firm f. Second, for each firm-day, we interpolate the previous day CPUE of sets by the firm's vessels using inverse distance weighting. Then for each set we extract lagged firm-level CPUE at that set's location. This procedure yields two variables from firms' lagged CPUE: the nearest value and the interpolated value.

Our third and fourth variables are firm-day demeaned nearest and interpolated lagged CPUE. Fifth, we record the distance between set i and the nearest set by vessels in the same firm on the previous day. Sixth, we calculate the average interpolated firm-level lagged CPUE in set i's best local zone. Our seventh and eighth variables are the differences between this sixth variable and set i's nearest and interpolated firm-level lagged CPUE.

Our ninth and tenth variables are the squared distances between the zone set i occurs in and the *predicted* best local zones, which are the zones within 126 km that have the highest average interpolated and nearest firm-level lagged CPUE. (By contrast, the best local zone for our sixth predictor variable is the one within 126 km that has the highest actual CPUE among all sets on day t.) Our eleventh predictor variable is the number of zones that vessels belonging to firm f fished in on day $t-1$. We add six more predictor variables by calculating third-order polynomials for the three (non-differenced) CPUE variables (the first, second, and sixth variables). We can calculate third-order instead of secord-order polynomials because unlike in our estimation of the precision of the public signal, we are not constrained by our server's RAM.

Finally, we interact all 17 of our private predictor variables to obtain a matrix of private signals with 153 variables (17 choose 2, plus the original 17 variables). If firm f had no sets on day $t-1$, then we record all predictor variables as missing, except for the number of zones that vessels belonging to firm f fished in on day $t-1$.

C.4 Cross-validation procedure used in estimating the variance of public and private signals

Our 10-fold cross-validation procedure iteratively divides the training set into an analysis set and an assessment set. One iteration involves training the model in an analysis set, consisting of 90% of days in the training set, and then assessing accuracy by predicting squared distance to the best local zone for the assessment set, consisting of the remaining 10% of days in the training set. We repeat the procedure 10 times, using different splits of the training data into analysis and assessment sets each time. We save the prediction accuracy—the R^2 in the assessment sets—at different values of the lasso penalty term. In each iteration of our cross-validation procedure, we impute missing predictor values with that variable's mean value in the analysis set, remove zero variance predictors (if any), and standardize all predictor variables. We define the optimal penalty term as in Breiman et al. (1984) as the penalty that returns the simplest model (fewest predictor variables) that is within one standard error of the numerically optimal penalty (the smaller penalty that results in a more complex model and the highest in-sample R^2 among all penalties).

C.5 Other specifications and robustness checks for estimating r

We considered several other dependent variables for our regressions, but we believe that squared distance to the best zone matches our model most closely. For example, we could choose CPUE of set i as our dependent variable, but predicting CPUE (which is not itself a location) is less similar to predicting the best location than our primary dependent variable.

Alternatively, we could estimate r with a two-step procedure in which our dependent variable equals 1 if the set occurs in the best local zone and equals 0 otherwise. In the first step, we would regress this variable on public or private signals and then use the resulting predicted values to obtain the predicted best local zone(s) each day. In the second step, we would calculate the precision of the public or private signal as the average squared distance between the predicted best and actual best local zone(s), where the predicted best local zone(s) comes from either the public signal regression or the private signal regression. This procedure is similar to our primary dependent variable, but we do not implement it because its two steps may degrade the precision of the signals relative to our primary one-step procedure.

We do, however, repeat our estimation procedure with an alternative measure of CPUE: tons per set minus vessel-level average tons per set. We obtain similar results compared to our primary specification of CPUE: r equals 0.905 (first row of Table C1), compared to 0.923 in our primary specification.

As a second robustness check, we repeat our estimation procedure after adjusting CPUE for travel costs (second row of Table C1). If anchoveta are slightly less abundant in Zone A than in Zone B, but Zone A is much closer to the firm's vessel, then Zone A may be the more profitable zone to fish in. Since we do not observe where each trip begins, we proxy for travel cost by calculating the distance between set i and the centroid of the sets by the vessel on its previous trip. We construct an alternative measure of CPUE by regressing tons caught on this distance to a new location and the firm and vessel characteristics in our primary specification. We obtain a similar parameter estimate as in our primary specification (r equals 0.910, compared to 0.923 in our primary specification).

Finally, we repeat our estimation using alternative radii to define the best local zone. This version changes our outcome variable, squared distance to the best local zone, as well as some of our public and private signals. The third row of Table C1 displays the results when we halve our preferred radius from 126 km to 63 km, and the fourth row displays the results when we double our preferred radius from 126 km to 252 km. In the half radius specification, the precision of the public signal declines relative to that of the private signal. The double radius specification reverses that relationship; the precision of the public signal increases relative to the precision of the private signal. Over larger areas, it seems that geophysical variables predict fishing grounds better than the previous day's CPUE by a firm's vessels. For this reason the BLUE of θ, the parameter δ, places more weight on the public signal when the area of feasible fishing grounds is larger.

We considered augmenting our private signal matrix with historical variables because past

Table C1: r robustness checks

Specification	Parameter estimates σ_y^2	σ_x^2	r	δ
Simple CPUE	8.82187e+11	1.07608e+12	0.905	0.550
Travel cost	9.29433e+11	1.12195e+12	0.910	0.547
Half radius	6.38564e+10	6.36043e+10	1.002	0.499
Double radius	1.56387e+13	3.13934e+13	0.706	0.667

Notes: Each row presents estimates from a different robustness check. We repeat our estimation procedure with a simpler measure of CPUE, tons per set minus vessel-level average tons per set, in the first row; with a measure of CPUE that accounts for travel costs in the second row; with a radius of 63 km, half that of our primary radius, in the third row; and with a radius of 252 km, double that of our primary radius, in the fourth row.

fishing success beyond the day-before could inform location decisions. For a given set i, we identify the vessel's best set in the previous year within 126 km of i. We calculate the distance from last year's best set to set i's best local zone, and we also record the CPUE of last year's best set. We add four new predictor variables to our private signal matrix: squared distance between last year's best set and i's best local zone, and a third-order polynomial of the CPUE of last year's best set. After interacting all predictor variables with each other, we re-estimate σ_x^2 with a matrix of 231 variables (21 choose 2, plus the original 21 variables). Estimating Equation 10 with our optimal shrinkage penalty (157) retains the same 8 variables with non-zero coefficients as in our primary specification without historical predictor variables. Our estimate of σ_x^2 is therefore the same (1.07994e+12). Historical CPUE may fail to predict squared distance to the best local zone because the spatial abundance of the anchoveta stock is not stable over time (Castillo et al., 2019).

C.6 Other specifications and robustness checks for estimating τ

Double-lasso procedure. Our application of the post-double selection method of Belloni, Chernozhukov, and Hansen (2014) in Section 7 has three steps (Column 2 of Table 2). First, we run a lasso regression with CPUE as the dependent variable (omitting standardized dispersion as a predictor variable). Second, we run a lasso regression with standardized dispersion as the dependent variable. Finally, we regress (via ordinary least squares) CPUE on standardized dispersion and include as controls all of the variables that lasso retains as predictors of CPUE or standardized dispersion. Standardized squared distance to the best local zone is one of the control variables in this final regression because it was retained as a predictor in both lasso regressions.

Table C2: τ robustness checks

	CPUE		
	(1)	(2)	(3)
Dispersion	1.585	0.450	1.587
	(0.810)	(0.623)	(0.956)
DistToBest^2	-4.223	-3.528	-3.670
	(0.757)	(0.848)	(0.910)
τ	0.375	0.128	0.432
	(0.211)	(0.179)	(0.255)
Adj. R^2	0.007	0.004	0.006
N	246,920	246,920	246,920

Notes: Each column presents estimates of Equation 12 from a different robustness check. We re-estimate Equation 12 with a simpler measure of CPUE, tons per set minus vessel-level average tons per set, in Column 1; with a measure of dispersion that accounts for the spatial configuration of sets in Column 2; and with a measure of CPUE that accounts for travel costs in Column 3. In all regressions, the dependent variable is CPUE and standard errors are two-way clustered at the date and zone level. In all columns, we estimate τ as the Dispersion coefficient divided by the negative of the DistToBest^2 coefficient, and we estimate the standard error of τ with the delta method.

In addition to including as predictors all of the public signal variables we used to estimate σ_y^2, we construct the following measures of local stock depletion, which is negatively correlated with contemporaneous biomass. For a given set i in zone z on day t, we calculate eight lags of tons caught in zone z: tons caught in zone z on day t prior to the start of set i, tons caught in zone z on day $t-1$, tons caught in zone z on day $t-2$, and so on until tons caught in zone z on day $t-7$. For each of these 8 variables, we calculate squared terms and indicators for 0 tons caught, yielding a matrix of 24 predictor variables. We add these 24 variables to the 180 public signal variables, and then interact all 204 variables with each other. We include calendar date indicator variables, zone indicator variables, and squared distance to the best local zone as predictors as well, but we do not include them in the interaction due to computational constraints. Zone indicator variables partially capture differences in the cost of fishing across locations. We thus obtain a matrix of 21,411 predictor variables. Of these, the two lasso regressions retain 265 unique predictor variables with non-zero coefficients.

Additional robustness checks. We consider three alternative specifications for estimating τ in Table C2. As in our primary specification, we always standardize dispersion and squared distance to the best local zone before re-estimating Equation 12 so that the two regression coefficients are directly comparable.

First, we re-estimate Equation 12 with our simpler measure of CPUE, vessel-demeaned tons per set. This robustness check returns estimates $B = 1.585$, $A = -4.223$, and $\tau = 0.375$ (Column 1 of Table C2). These are similar to the estimates in our primary CPUE specification (Column 1 of Table 2).

Second, we consider a dispersion measure that accounts for sets' spatial configuration:[25]

$$D'_{it} = \frac{1}{M_{it}-1} \cdot \frac{\sum_{j\neq i}(k_{it}-k_{jt})^2}{\left(\frac{\#\text{ sets in NE}}{M_{it}-1}\right)^2 + \left(\frac{\#\text{ sets in SE}}{M_{it}-1}\right)^2 + \left(\frac{\#\text{ sets in SW}}{M_{it}-1}\right)^2 + \left(\frac{\#\text{ sets in NW}}{M_{it}-1}\right)^2} \quad (47)$$

where D'_{it} is the alternative measure of dispersion, $M_{it}-1$ is the number of other sets within 126 km of set i, $\sum_{j\neq i}(k_{it}-k_{jt})^2$ is the sum of squared distances between set i and sets j among sets j that are within 126 km of set i; and NE, SE, SW, and NW denote the northeast, southeast, southwest, and northwest quadrants of the circle with radius of 126 km and center at k_{it}. For example, if all sets j are in i's northeast quadrant, they could have a smaller effect on i's CPUE than if the j sets were more evenly distributed across quadrants, since in the latter case vessels have arguably dispersed to a greater extent. Accounting for the spatial configuration of sets could therefore increase the estimated effect of dispersion on CPUE, which would likely increase our estimate of τ. Our reasoning is, if D'_{it} provides a more accurate measure of dispersion, then our use of D_{it} in our primary specification introduces measurement error. Although the effect of measurement error on coefficient estimates is ambiguous in general, it often biases estimates toward zero. Based on this logic, we would expect that replacing D_{it} with D'_{it} would lead to a larger (in absolute value) coefficient estimate if D'_{it} were actually a more precise measure of dispersion. However, when we re-estimate Equation 12 with D'_{it} as our measure of dispersion, we estimate a smaller effect of dispersion and a smaller τ (Column 2 of Table C2). Therefore, we have no grounds for thinking that D'_{it} provides a more accurate measure of dispersion.

Third, Column 3 of Table C2 presents an estimate of τ using our measure of CPUE that accounts for travel costs (Online Appendix C.5). We obtain a nearly identical estimate, 0.432, compared to the estimate of τ in our primary specification (0.448).

References

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, *81*(2), 608–650.

[25] We thank an anonymous referee for suggesting this specification.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth.

Castillo, R., Rosa, L. D., Diaz, W. G., Madureira, L., Gutierrez, M., Vásquez, L., & Koppelmann, R. (2019). Anchovy distribution off Peru in relation to abiotic parameters: A 32-year time series from 1985 to 2017. *Fisheries Oceanography*, *28*(4), 389–401.

Flanders Marine Institute. (2019). Exclusive Economic Zones (200NM), version 11 [https://doi.org/10.14284/386].

NASA. (2014). MODIS-Aqua Chlorophyll-a Data (Level 3, mapped, daily, 4 km) [NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Accessed on April 21, 2021]. https://doi.org/10.5067/AQUA/MODIS_OC.2014.0

NASA. (2016a). VIIRS-SNPP Sea Surface Temperature Data (Level 3, mapped, daily, 4 km, 11 µm daytime) [NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Accessed on January 21, 2022]. https://doi.org/10.5067/SNPP/VIIRS/L3B/SST/2016

NASA. (2016b). VIIRS-SNPP Sea Surface Temperature Data (Level 3, mapped, daily, 4 km, 11 µm nighttime) [NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Accessed on January 21, 2022]. https://doi.org/10.5067/SNPP/VIIRS/L3B/SST/2016

NASA. (2018). VIIRS-SNPP Chlorophyll-a Data (Level 3, mapped, daily, 4 km) [NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Accessed on January 21, 2022]. https://doi.org/10.5067/SNPP/VIIRS/L3M/CHL/2018

NASA. (2019a). MODIS-Aqua Sea Surface Temperature Data (Level 3, mapped, daily, 4 km, 11 µm daytime) [NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Accessed on April 21, 2021]. https://doi.org/10.5067/MODSA-1D4D9

NASA. (2019b). MODIS-Aqua Sea Surface Temperature Data (Level 3, mapped, daily, 4 km, 11 µm nighttime) [NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Accessed on January 21, 2022]. https://doi.org/10.5067/MODSA-1D4N9

NOAA. (2021). SMAP Sea Surface Salinity Data (Daily, 0.25° resolution) [NOAA CoastWatch. Accessed on May 10, 2021. https://www.star.nesdis.noaa.gov/data/socd1/coastwatch/products/miras/nc/].

NOAA. (2022). Sea Surface Temperature Anomaly Data (Daily, 25 km resolution) [Accessed on January 21, 2022]. http://ihma.org.pe/anomalia-termica-de-la-superficie-del-mar/

NOAA. (2023). Sea Level Anomaly and Geostrophic Currents Data (Daily, 0.25° resolution) [NOAA Laboratory for Satellite Altimetry. Accessed on August 21, 2023. https://coastwatch.noaa.gov/cwn/products/sea-level-anomaly-and-geostrophic-currents-multi-mission-global-optimal-interpolation.html].