Does more Information-gathering Effort Raise or Lower the Average Quantity Produced?

Thomas Marschak*  J. George Shanthikumar†  Junjie Zhou‡

January 31, 2017

Abstract

We aim at some simple theoretical underpinnings for a complex empirical question studied by labor economists and others: does Information-technology improvement lead to occupational shifts — toward “information workers” and away from other occupations — and to changes in the productivity of non-information workers? In our simple model there is a Producer, whose payoff depends on a production quantity and an unknown state of the world, and an Information-gatherer (IG) who expends effort to learn more about the unknown state and then sends the Producer a signal. The Producer responds by revising prior beliefs about the states and using the posterior to make an expected-payoff-maximizing quantity choice. We consider a variety of IGs and variety of Producers. For each IG there is a natural effort measure. Our central aim is to find conditions under which more IG effort leads to a larger average production quantity (“Complements”) and conditions under which it leads to a smaller average quantity (“Substitutes”). We start by considering Blackwell IGs, who meet the strong conditions required in the Blackwell theorems about the comparison of experiments. We then turn to non-Blackwell IGs, where the Blackwell theorems cannot be used and special techniques are needed to obtain Complements/Substitutes results.

Keywords: Information technology and productivity, Blackwell Theorems, Garbling

JEL classifications: C44, D24, L11.

---

*Walter A. Haas School of Business, University of California, Berkeley. e-mail: marschak@berkeley.edu
†Krammert School of Management, Purdue University. e-mail: jshanthi@purdue.edu
‡Department of Economics, National University of Singapore. e-mail: zhoujj03001@gmail.com
1. Introduction.

This paper aims to contribute some simple but new theoretical underpinnings to the study of a complex but empirically well-motivated question: does more informational effort lead producers to produce more or to produce less, to use more “non-information workers” or fewer? The workforce-composition question is an ongoing challenge for labor economists. A prominent approach identifies “cognitive” and “routine” tasks and studies an employer who hires the workers that perform them.\(^1\)

To connect the “task” approach to the problem we study here, consider the following highly simplified interpretation of that approach. A firm employs the quantity \(L_1\) of routine-task labor and \(L_2\) of cognitive-task labor. The firm uses two types of capital, IT-related capital and non-IT capital; their quantities are \(K_1\) and \(K_2\). IT investment improves the productivity of the cognitive labor and reduces the amount of noncognitive labor needed for a given output. The firm’s product quantity is described by a production function, say \(\phi(L_1, L_2, K_1, K_2)\). The firm obtains the revenue \(R(t)\) for the product quantity \(t\). It chooses the four input quantities so as to maximize \(R((\phi(L_1, L_2, K_1, K_2)) - \text{the cost } w_1L_1 + w_2L_2 + w_1^*K_1 + w_2^*K_2\), where \(w_1, w_2, w_1^*, w_2^*\) are factor prices. The effect of changes in the factor prices on the two labor quantities can be studied, as well as the effect of a shift in the production function itself as a result of technology improvement.

Our model is much simpler. It contains just two persons. There is a Producer who has to choose a production quantity and the value of his choice depends on an unknown state variable with a known prior. The Producer uses an Information Gatherer (IG) who studies the state variable and then sends a signal to the Producer, who revises the prior and chooses a production quantity that maximizes expected payoff under the posterior. The IG the Producer uses is chosen from a variety of IGs. Each of them performs some information-gathering service — sampling, polling, forecasting, inspecting, interviewing, “data mining” — which is useful to many different Producers. We ask whether more IG effort leads to a rise or a drop in the average quantity the Producer chooses. If more IG effort leads the Producer to increase (decrease) average quantity, and if the non-information labor that the Producer needs rises linearly when his chosen quantity rises, then when we see an increase in the IG’s workforce (because IG effort has increased) we will also see an increase (decrease) in the Producer’s average non-information workforce. If the required non-information labor rises nonlinearly when product quantity increases, then further study is needed.

\(^1\)A seminal contribution is Autor, Levy, and Murnane (2003). They are interested in the effects of increased computer investment on the hired quantities of routine-task labor and cognitive-task labor. A recent extensive discussion of the task identification problem is found in Acemoglu and Autor (2011). See also Machin and Reenen (1998) and Wolff (2002).
It will be convenient to let Complements denote the case where more IG effort leads to higher or unchanged average production quantity and to let Substitutes denote the case where more IG effort leads to lower or unchanged average production quantity. The IG/Producer model covers a wide variety of situations\(^2\) and we obtain a varied collection of Complements/Substitutes results.

2. The model.

The Producer chooses the product quantity \(q\). The choice yields a payoff \(u(q, \theta)\). Here \(\theta\) is a random state variable. Its possible values comprise the set \(\Theta\). The Producer has a prior on \(\Theta\). The IG studies the states and has a collection of available experiments on the set \(\Theta\). An experiment is defined by a signal set \(Y\) and a set of likelihood distributions on \(Y\), one for every state in \(\Theta\). Once the Producer receives a signal from the IG, he replaces his prior with a posterior and chooses a quantity that maximizes the expected value of \(u\) under that posterior.

The IG’s available experiments differ with regard to some natural measure of effort. Note that a prior on the states is not specified when we construct an experiment. If we do specify a prior, then the experiment defines what we shall call an information structure. A structure consists of (1) a (marginal) probability distribution on the IG’s set of signals, and (2) a set of posteriors on the state set, one for each signal. For a given prior, we shall say that the first of two structures requires more effort than the second if the experiment which defines the first structure requires more effort than the experiment which defines the second structure. An IG might, for example, be a Partitioning Refiner. He partitions the possible values of \(\theta\) and his signal tells the Producer the set in which \(\theta\) lies. At higher IG effort the partitioning is refined. The IG might be a One-point Forecaster. His signal gives the Producer a one-point forecast of the upcoming value of \(\theta\), and higher IG effort means an increase in the probability that the forecast is correct. The IG might

\(^2\)Here are a few of them. The Producer might be a monopolist who uses an IG to sample potential buyers and finds the prices they are willing to pay. More IG effort means a larger sample. After receiving the IG’s report, the Producer replaces his prior on the proportion of eager and “soft” buyers with a posterior and chooses a price that maximizes expected profit under the posterior. When the IG exerts more effort, does the average product quantity sold by the Producer rise or fall? The Producer might be an expected-profit-maximizing inventory manager who sells product at a fixed and known price, but has to order before knowing what the next period’s demand will be. The IG forecasts demand and more IG effort makes the forecast more reliable, in a precise sense, and more useful to the Producer. When the IG exerts more effort, does the Producer, on the average, order more or less? The Producer might be a lender who is uncertain about borrowers’ qualifications. The IG probes the qualifications, computes a credit score for each borrower, and sends the score to the Producer. More IG effort means that the reported score is, in an appropriate sense, a more reliable indicator of the borrower’s future behavior and is more useful to the Producer. When the IG exerts more effort, does the Producer, on the average, lend more or less? Going much further afield, the Producer might be a surgeon who performs a procedure following a certain diagnostic result. The IG is a diagnostican. By exerting more effort he can reduce the frequency of false positives. When the IG exerts more effort, does the Producer, on the average, perform more or fewer procedures?
be a Sampler, who samples a population. The Producer uses the sample to update his estimate of a population parameter. The parameter determines expected profit for each of the Producer’s possible quantities. Size of the sample is a natural IG-effort measure. We shall be studying these IGs and others as well.

We now complete the model and make it more precise. For a given experiment and a given state \( \theta \), we let \( \lambda_\theta \) denote the experiment’s likelihood CDF (or measure) on the experiment’s signal set \( Y \). Thus an experiment on the state set \( \Theta \) is a pair \( E = (Y, \{\lambda_\theta\}_{\theta \in \Theta}) \). When \( Y \) is finite, we sometimes abuse the “\( \lambda_\theta \)” notation and we let \( \lambda_\theta \) denote a vector of signal probabilities.

If we specify a prior CDF (or measure), say \( G \), then \( G \), together with the experiment \( E \), determine:

- \( F_y^G \), a posterior CDF (measure) on \( \Theta \), for every signal \( y \in Y \).
- \( W_y^G \), a marginal CDF (measure) on the signal set \( Y \).

The prior \( G \) and the IG experiment \( E \) determine the information structure \( I = (\{F_y^G\}_{y \in Y}, W_y^G) \). If the prior \( G \) is understood, we can omit the superscript \( G \). So once \( G \) is fixed, we may view the IG as having a collection of possible structures rather than a collection of possible experiments. For a given IG, we will consider a real-valued effort measure \( \eta \), defined on the IG’s collection of experiments and his collection of structures. We let \( \eta(E) \) denote the effort of the experiment \( E \), and (slightly abusing the “\( \eta \)” notation) we let \( \eta(I) \) denote the effort of the structure \( I \) defined by \( E \) and the prior. If the first of two structures has effort \( \eta^* \), the second has effort \( \eta^{**} \), and \( \eta^{**} > \eta^* \), we shall say that the IG works harder when he moves from the first structure to the second.

Now we turn to the Producer. He has prior beliefs \( G \) about the states. He also has a set \( A \subseteq \mathbb{R} \) of possible actions (production quantities) and a payoff function \( u : A \times \Theta \to \mathbb{R} \). We use the symbol \( q \) for the elements of \( A \). Here is a key to our terminology and notation:

- We shall say that the Producer’s payoff function \( u \) is regular for a given CDF (measure) on \( \Theta \), say \( F \), if there exists a largest maximizer of the expected value of \( u \) on the action set \( A \) under that CDF.
- The symbol \( \hat{q}(F; u) \) denotes that largest maximizer. For convenience, we frequently call the quantity \( \hat{q}(F; u) \) the Producer’s best quantity for the CDF (measure) \( F \) and the payoff function \( u \).

---

3We use two alternative symbols for expectation. If \( T \) is a CDF (cumulative density function), or a probability measure, then \( E_T \) denotes expectation under \( T \). If a random variable \( x \) has support \( X \), and a CDF or measure is understood, then \( E_{x \in X} \) will be an alternative symbol for expectation.
• We let $V_u(F)$ denote $E_F u(\hat{q}(F;u),\theta)$ (the expected value, under $F$, of the maximal $u$).

• For the IG/Producer pairs that we study we shall assume that the Producer’s payoff function $u$ is regular for all of the IG’s structures. That means that for every structure $(\{F_y\}_{y \in Y}, W_Y)$ the function $u$ is regular for every posterior $F_y$.

• Given the IG’s signal $y$, the Producer chooses the action $\hat{q}(F_y;u)$ (the best quantity for the posterior $F_y$).

• For the payoff function $u$, an experiment $E$, and the information structure $I$ which is defined by $E$ and the prior, we have a value of the experiment, namely the average, over all of the possible posteriors (signals), of the maximal expected value of $u$ under the posterior.

• If we want to consider several priors $G$, then we let the symbols $V_u^G(E)$ and $V_u^G(I)$ denote value.

We now have the concepts needed for a formal definition of Complements and Substitutes. Suppose we are given a prior $G$. Consider a Producer with a payoff function $u$. Consider an IG with a collection of possible experiments, each determining a structure. Assume that the payoff function $u$ is regular for all the posteriors in all the structures. Consider a real-valued effort measure $\eta$ on the collection of structures.

**Definition 1** We have Complements (Substitutes) if the following holds whenever the effort measure $\eta$ is higher for the structure $I' = (\{F'_y\}_{y \in Y'}, W_{Y'})$ than for the structure $I = (\{F_y\}_{y \in Y}, W_Y)$: the expected value of $\hat{q}(F'_y;u)$, over all the signals $y \in Y'$ of the structure $I'$, is not less than (not more than) the expected value of $\hat{q}(F_y;u)$, over all the signals $y \in Y$ of the structure $I$.

Informally: we have Complements (Substitutes) if the Producer’s average best quantity does not drop (does not rise) when the IG works harder.

Next we turn to Blackwell IGs.

**Definition 2** An IG is a Blackwell IG if there exists an effort measure $\eta$ on the IG’s collection of experiments such that for every prior $G$ and every regular payoff function $u$,

$$V_u^G(E') \geq V_u^G(E) \text{ whenever } \eta(E') > \eta(E).$$
Informally: *value never drops when effort rises.* It will sometimes be useful to say that a Blackwell IG has an effort measure with the *informativeness property.*

It is obvious that the Partitioning Refiner and the Sampler are Blackwell IG’s. Refining the partitioning can never hurt the Producer, no matter what the prior and the payoff function might be, since the Producer is free to ignore the refinement. If the refinement splits some set into two, for example, the Producer can always achieve his pre-refinement expected payoff by continuing to treat the two new sets as though they were one. If the sample size increases, the Producer can achieve his previous expected payoff by ignoring the additional observations. It is not obvious, however, that the One-point Forecaster is a Blackwell IG. It turns out that the One-point Forecaster is a special case of an IG whom we call the Combiner, studied in detail in Section 5 below. This IG has two “anchor” structures. The second anchor is more informative than the first. Every structure in the IG’s collection is a weighted combination of the two anchors and IG effort is measured by the weight given to the second anchor. A careful argument is needed to establish that the Combiner (and hence the One-point Forecaster) is, in fact, a Blackwell IG.

The Blackwell IG is important for the Complements/Substitutes question. To see this, we now consider a Finite Blackwell Theorem. The finite theorem asserts the equivalence of three statements, which we informally call “Informativeness”, “Garbling”, and “Convex Functions”. As we shall see, the “Convex Functions” statement immediately yields a Complements/Substitutes result whenever the Producer’s best quantity is a convex or concave function of the posterior. Our Finite Blackwell theorem is the one that is most familiar to economists. We shall consider a non-finite version of the theorem in Section 5.1.9 below.

A Finite Blackwell Theorem

Suppose that the state set is $\Theta = \{\theta_1, \ldots, \theta_i, \ldots, \theta_n\}$. Consider the experiment $\mathcal{E} = (Y, \{\lambda_\theta\}_{\theta \in \Theta})$, where $Y = \{y_1, \ldots, y_m\}$ and $\lambda_\theta$ denotes the likelihood vector $(\lambda_1^\theta, \ldots, \lambda_j^\theta, \ldots, \lambda_m^\theta)$. Consider also the experiment $\mathcal{E}' = (Y', \{\lambda'_\theta\}_{\theta \in \Theta})$, where $Y' = \{y'_1, \ldots, y'_m\}$ and $\lambda'_\theta$ denotes the likelihood vector $(\lambda^n_1^\theta, \ldots, \lambda^n_j^\theta, \ldots, \lambda^n_m^\theta)$. The following three statements are equivalent:

(a) (“Informativeness”) For every payoff function $u : A \times \Theta \to \mathbb{R}$, and for every prior $G$ such that $u$ is regular for all of the posteriors that are identified by the signals of both experiments, we have $V_u^G(\mathcal{E}') \geq V_u^G(\mathcal{E})$.

---

4The two fundamental Blackwell papers are Blackwell (1951) and (1953). Proofs, addressed to an economic audience, of the equivalence of “Informativeness” and “Garbling” when state and signal sets are finite, are found in McGuire (1972), Ponsard (1975), Crémer (1982), and Leshno and Spector (1992). An extensive treatment of the problem of comparing Information Gatherers when the state and signal sets are finite, including the “Convex Functions” statement of the finite Blackwell theorem, is given in Marschak and Miyasawa (1968).
(b) ("Garbling") Let \( \Lambda_{ij} \) denote \( \lambda^{i \downarrow}_{b_{ij}} \) and let \( \Lambda'_{ij} \) denote \( \lambda^{i \downarrow}_{b_{ij}} \). Consider the two likelihood matrices \( \Lambda = \left( \left( \Lambda_{ij} \right) \right) \) and \( \Lambda' = \left( \left( \Lambda'_{ij} \right) \right) \). There exists a row-stochastic \( m' \)-by-\( m \) "garbling" matrix \( B \) such that \( \Lambda \overset{n \text{-by-} m}{\sim} \Lambda' \overset{n \text{-by-} m'}{\sim} B \).

(c) ("Convex Functions") For experiment \( \mathcal{E} \) and a given prior, let \( F_y \) denote the posterior identified by the signal \( y \in Y \). For experiment \( \mathcal{E}' \) and the same prior, let \( F'_{y'} \) denote the posterior identified by the signal \( y \in Y' \). Let \( \phi \) be any real-valued convex function on \( \Delta(\Theta) \). Then \( E_{y \in Y'} \phi(F'_{y'}) \geq E_{y \in Y} \phi(F_y) \).

Even though the inequality in Statement (a) is weak, it will be convenient to summarize it informally by saying that \( \mathcal{E}' \) is more informative than \( \mathcal{E} \). The garbling matrix \( B \) in statement (b) shows, in a precise manner, how information about the states is degraded when we move from the experiment \( \mathcal{E}' \) to the experiment \( \mathcal{E} \). Statement (c), whose equivalence to (a) we shall often use, has not been as prominent in economic discussions as statements (a) and (b). The proof that (c) is equivalent to (a) uses the garbling matrix of statement (b) and is straightforward.  

---

5Given a prior on the states, one considers experiment \( \mathcal{E}' \)'s \( n \)-by-\( m' \) matrix \( \Pi' = \left( (\pi'_{ij}) \right) \) of posterior probabilities, where \( \pi'_{ij} \) is the probability of state \( \theta_i \) given the \( \mathcal{E}' \) signal \( y'_{ij} \), and the analogous \( n \)-by-\( m \) matrix \( \Pi = \left( (\pi_{ij}) \right) \) for experiment \( \mathcal{E} \). Using the garbling matrix of statement (b) one constructs another garbling matrix, this time for the posteriors. One is denoted \( \overline{B} \); its typical element is \( \overline{b}_{jk} \); it is \( m' \)-by-\( m \) and column-stochastic; and it satisfies: \( \Pi = \Pi' \cdot \overline{B} \). Let \( \overline{\pi}^j_k \) denote two members of \( \Delta(\Theta) \), namely the vector of posterior state probabilities for \( y'_{ij} \) (one of \( \mathcal{E}' \)'s \( m' \) signals), and for \( y_{ij} \) (one of \( \mathcal{E} \)'s \( m \) signals). We find that for every \( k \in \{1, \ldots, m\} \), we have

\[
\overline{\pi}^j_k = \sum_{j \in \{1, \ldots, m'\}} \overline{\pi}^j_{jk} \cdot \overline{b}_{jk}.
\]

We also find that for the signal probabilities we have

\[
\Pr(y'_{ij}) = \sum_{k \in \{1, \ldots, m\}} \overline{b}_{jk} \cdot \Pr(y_{ij})
\]

for every \( j \in \{1, \ldots, m'\} \). Given a convex function \( \phi \) on \( \Delta(\Theta) \), we then have (using Jensen’s inequality):

\[
\sum_{j \in \{1, \ldots, m'\}} \Pr(y'_{ij}) \cdot \phi(\overline{\pi}^j_{ij}) \geq \sum_{k \in \{1, \ldots, m\}} \Pr(y_{ij}) \cdot \phi(\overline{\pi}^j_k).
\]

Thus statement (b) implies statement (c). To show that (c) implies (a), we use the well-known fact (reviewed below in section 4.1.2) that for any CDF \( F \) on \( \Theta \) the highest attainable value of \( E_{F(u)}(q, \theta) \) (which we have denoted \( V_a(F) \)) is a convex function of \( F \).

6In Marschak, Shanthikumar and Zhou (2014) we introduce the concept of a strict garbling matrix, which satisfies several weak conditions in addition to row-stochasticity. We show that if \( \Lambda', \Lambda \) satisfy (b) for a strict matrix \( B \), then a new version of (c) holds. In the new version the final inequality is strict whenever \( \phi \) is strictly convex. That allows us to obtain strict Complements/Substitutes results. If, for example, the Producer’s payoff function is \( u(q, \theta) = \theta q - \frac{1}{1+q} q^{1+k} \) (where \( 0 < k < 1 \)), then best quantity is a strictly convex function of the posterior mean. If the Blackwell IG whom this Producer uses moves from \( \Lambda \) to the higher-effort \( \Lambda' \) (where \( B \) is strict), then the Producer’s average best quantity, over all the posteriors, strictly increases.
If the Producer’s best quantity turns out to be a convex or concave function of the posterior, and if our effort measure for the IG has the informativeness property (i.e., the IG is a Blackwell IG), then the “Convex Functions” statement gives us a Complements or Substitutes result: higher IG effort leads to a rise in the Producer’s average best quantity in the convex case and a drop in the concave case. If average best quantity is neither convex nor concave in the posterior, then we cannot use the “Convex Functions” statement. Fresh study is required to obtain Complements/Substitutes results, even if the Producer uses a Blackwell IG.

3. Related literature.

We can partially order the experiments in a Blackwell IG’s collection of possible experiments so that a higher-ranked experiment is more useful to any Bayesian decision-maker who responds to each of the experiment’s signals by maximizing the expected utility of some von Neumann-Morgenstern utility function (our $u(q, \theta)$). If the collection is large, there may be many experiment pairs that are non-comparable. That is troublesome in many applications. Accordingly, a number of papers have restricted the utility indicator. For example, Lehmann (1988) and Persico (2000) consider utility indices that satisfying the single-crossing property, Athey and Levin (2001) study supermodular utility indices, and Quah and Strulovici (2009) explore interval-dominance-order utilities. This literature, however, does not deal directly with our main concern — the average quantity chosen in response to an experiment’s signals.

In many of our applications only the mean of the posterior matters to the decision maker. A recent literature considers such decision makers and ranks information structures using various dispersion orderings of conditional expectations. The contributions include Gauzuza and Penalva (2010) and Brandt, Drees, Eckwort, and Vardy (2013). The least variable structure conveys no information beyond the prior, and the most variable structure conveys perfect information. So one might interpret variability as an effort measure. Under certain conditions a higher-effort structure is more informative in the Blackwell sense. That may imply that average quantity rises (falls) when effort increases, but the question is not pursued in these papers.

Kamenica and Gentzkow (2011) study a model of Bayesian persuasion. The sender’s signal induces an action by the receiver and the sender has preferences over the possible actions. The paper considers the set of information structures which are “Bayes plausible” (expected posterior probability equals the prior) and finds the signal that is optimal from the sender’s perspective. One could make a connection to our problem if the receiver’s action is our “quantity” and the sender prefers the higher (lower) of two quantities. The paper, however, solves for the optimal signal rather than comparing two entire experiments.

Several papers are related to the IG whom we call the One-point Forecaster. The papers
consider an IG who either conveys perfect information to a decision-maker or else uninformative noise. The probability of the perfect-information signal can be varied. Examples include Lewis and Sappington (1992, 1994) and Johnson and Myatt (2006). We generalize the One-point Forecaster to the much wider class of IGs whom we call Combiners.

The design literature contains a great many papers on incentive problems in which a self-interested person acquires information and may convey it to others. The information gatherer may be a bidder in an auction (Bergemann and Pesendorfer (2007), Shi (2012), and many others) or the agent in a principal/agent problem (Szalay (2008), Krishna and Morgan (2008), and many others). The effect of more information-gathering on quantity is not directly studied in these papers.

In summary, the effect of more IG effort on average quantity is rarely the direct concern of the many papers that vary the information available to economic deciders. The problem we study here seems to have escaped intensive scrutiny.

4. Plan of the remainder of the paper.

In Section 5 we consider two specific Blackwell IGs: the Combiner and the Erratic IG. Section 6 considers two non-Blackwell IGs, the Scale/Location Transformer and the Equal-probability Partitioner. Section 7 presents some economic applications, where the Producer is a monopolist or price-taker and the IG is a Blackwell IG. Concluding remarks are made in Section 8. The proofs are relegated to the appendix.

5. Two specific Blackwell IGs: the Combiner and the Erratic IG

5.1 The IG who combines two anchor structures: the finite case.

For brevity we shall call this IG the Combiner. We consider first the case where both the state set and the signal set are finite. Once the prior is fixed, the Combiner has a family of structures $I_k$, where $k$ takes all the values in $[0, 1]$. All structures in the family have the same signal set. The structure $I_k$ combines two anchor structures, an initial structure $I_0$ and a more informative terminal structure $I_1$. In the structure $I_k$, the likelihood matrix equals $1 - k$ times the likelihood matrix of $I_0$ plus $k$ times the likelihood matrix of $I_1$.

\footnote{An early literature compares a monopolist’s decisions when he is uncertain about the demand curve with his decisions when he knows the demand curve with certainty (Sandmo (1971), Leland (1972), Holthausen (1976)). If the monopolist is risk-averse, the quantity chosen under certainty is higher than the quantity chosen under uncertainty. That can be interpreted as a rather primitive “Complements” result: if the monopolist uses an IG to acquire information about demand, and the IG moves from an effort that yields less-than-perfect information to an effort that yields perfect information, then average quantity rises.}
A natural effort measure for a Combiner is \( k \), the weight he places on the more informative anchor structure. One might, for example, interpret \( k \) as the probability that the Combiner uses \( I_1 \) and \( 1 - k \) as the probability that he uses \( I_0 \).

**Definition 3 (Null structures and Perfect structures)** An anchor structure is **null** whenever all the rows of its likelihood matrix are identical. An anchor structure is **perfect** if each signal uniquely identifies the true state.

It might be the case that \( I_0 \) is null and \( I_1 \) is perfect. As the Combiner expends more effort, we move further from the useless null case (where the posterior is the same for all signals) and closer to the perfect case. But our results go beyond the case where one anchor is null and the other is perfect. We shall study the more general case where the initial anchor may be more informative than a null structure and the final anchor may be less than perfect.

Not all Combiners are Blackwell IGs. We shall be particularly interested in those who are. For a Blackwell Combiner the effort measure \( k \) has the informativeness property: the structure \( I_{k'} \) is more informative than the structure \( I_k \) (or, equivalently, \( I_k \) is a garbling of \( I_{k'} \)) whenever \( 0 \leq k < k' \leq 1 \). Value never drops when \( k \) rises. We show the following:

- A Combiner is a Blackwell IG if and only if the following holds: for any regular payoff function and any prior, value is minimized at the initial anchor \( I_0 \).
- A Combiner for whom \( I_0 \) is null must be a Blackwell IG.
- There are Combiners who are Blackwell IGs even though \( I_0 \) is not null.
- There are Combiners for whom \( I_0 \) is not null who are not Blackwell IGs.
- For any non-null \( I_0 \) we can find a terminal anchor \( I_1 \) such that the Combiner is not a Blackwell IG.

We shall establish the first four of these five statements. A proof of the fifth statement is found in Marschak, Shanthikumar, and Zhou (2014).

Formally, let the state set be \( \Theta = \{ \theta_1, \ldots, \theta_n \} \subset \mathbb{R} \). We define the combiner IG as follows:

**Definition 4 (The Combiner IG)** Let \( \Lambda^1 \) and \( \Lambda^0 \) be the likelihood matrix of two anchors \( I_1 \) and \( I_0 \) with the same number of signals, namely \( m \). For each \( k \in [0, 1] \), the combiner structure \( I_k \) uses the experiment \( E_k \), which has \( m \) signals, denoted \( y_1^k, \ldots, y_m^k \) with likelihood matrix given by

\[
\Lambda^k = ((\Lambda^k_{ij})) = k \cdot \Lambda^1 + (1 - k) \cdot \Lambda^0.
\]
If the prior state probabilities are \( g_1, \ldots, g_n \), then we have the following matrix of joint signal/state probabilities:

\[
\begin{pmatrix}
\sum_{i=1}^{n} g_i \cdot [k \Lambda_{i1}^1 + (1-k) \cdot \Lambda_{i0}^1] & \sum_{i=1}^{n} g_i \cdot [k \Lambda_{i2}^1 + (1-k) \cdot \Lambda_{i0}^1] & \cdots & \sum_{i=1}^{n} g_i \cdot [k \Lambda_{im}^1 + (1-k) \cdot \Lambda_{i0}^1] \\
\sum_{i=1}^{n} g_i \cdot [k \Lambda_{i1}^2 + (1-k) \cdot \Lambda_{i0}^2] & \sum_{i=1}^{n} g_i \cdot [k \Lambda_{i2}^2 + (1-k) \cdot \Lambda_{i0}^2] & \cdots & \sum_{i=1}^{n} g_i \cdot [k \Lambda_{im}^2 + (1-k) \cdot \Lambda_{i0}^2] \\
\sum_{i=1}^{n} g_i \cdot [k \Lambda_{i1}^3 + (1-k) \cdot \Lambda_{i0}^3] & \sum_{i=1}^{n} g_i \cdot [k \Lambda_{i2}^3 + (1-k) \cdot \Lambda_{i0}^3] & \cdots & \sum_{i=1}^{n} g_i \cdot [k \Lambda_{im}^3 + (1-k) \cdot \Lambda_{i0}^3]
\end{pmatrix}
\]

If we delete all the \( g_i \) terms, we have experiment \( E_k \)'s likelihood matrix \( \Lambda^k \).

5.1.1 A lemma about the relation between \( k \) and any convex function of \( I_k \)'s posteriors.

We now state a Lemma which will be a useful tool. It concerns any anchor pair \((I_0, I_1)\), all the Combiner’s structures \( I_k \), and any convex (concave) function of the posteriors defined by \( I_k \)'s signals. For the signal \( y^k \), the posterior probability of the state \( \theta_i \) is \( u_{ij}(k) \equiv g_i [k \Lambda_{ij}^1 + (1-k) \cdot \Lambda_{ij}^0] \). The (marginal) probability of the signal \( y^k \) is \( w^k \equiv \sum_{i=1}^{n} g_i [k \cdot \Lambda_{ij}^1 + (1-k) \cdot \Lambda_{ij}^0] \). The posterior probability vector \((u_{ij}(k), \ldots, u_{nj}(k))/w^k\) belongs to the probability simplex \( \Delta(\Theta) = \{(z_1, \ldots, z_n) : z_i \geq 0, \forall i; z_1 + \cdots + z_n = 1\} \).

Consider a function \( \chi : \Delta(\Theta) \to IR \) and let \( t_\chi(k) \) denote the average value of \( \chi \) over the \( m \) posterior-probability vectors of the structure \( I_k \). Thus

\[
t_\chi(k) = \sum_{j=1}^{m} w_j^k \cdot \chi \left( \frac{u_{1j}(k)}{w^k_j}, \ldots, \frac{u_{nj}(k)}{w^k_j} \right).
\]

Lemma 1

If \( \chi : \Delta(\Theta) \to IR \) is convex (concave), then \( t_\chi : [0,1] \to IR \) is also convex (concave). If \( t_\chi \) is maximized (maximized) at \( k = 0 \), then \( t_\chi \) is nondecreasing (nonincreasing) if \( t_\chi \) is convex (concave).

The proof (in Appendix A) uses the fact that both \( u_{ij}(k) \) and \( w^k_j \) are linear in \( k \).

5.1.2 A further Lemma and a theorem about Combiners who are Blackwell IGs.

Recall that for any \( F \) in \( \Delta(\Theta) \) we say that the payoff function \( u \) is regular for \( F \) if there exists a quantity \( \hat{q}(F; u) \) which is a largest maximizer of \( E_F \ u(q, \theta) \). We call it the best quantity for the pair \((F, u)\). Then \( V_u(F) \) denotes the expected payoff, under the distribution \( F \), when that best...
quantity is used. We shall need a well-known general fact about the function $V_u : \Delta(\Theta) \to \mathbb{R}$: the function is convex.\footnote{Consider two distributions on $\Theta$, say $L$ and $M$, and consider $\lambda \in [0,1]$. Suppose that $q^*$ is a maximizer of $E_{\lambda L + (1-\lambda) M} u(q, \theta)$. Then:

$$V_u(\lambda L + (1-\lambda) M) = E_{\lambda L + (1-\lambda) M} u(q^*, \theta) = \lambda E_L u(q^*, \theta) + (1-\lambda) \cdot E_M u(q^*, \theta) \leq \lambda \max_q E_L u(q, \theta) + (1-\lambda) \cdot \max_q E_M u(q, \theta) = \lambda V_u(L) + (1-\lambda)V_u(M).$$}

Recall that $V_u(I)$ denotes the value of the structure $I$ for the payoff function $u$. We let $V_u(k)$ be an abbreviation for $V_u(I_k)$. It will be convenient to introduce a new symbol $Q_u(I)$ when $u$ is regular for $I$. The symbol denotes the expected value of the best quantity, over all of $I$’s posteriors. We abbreviate $Q_u(I_k)$ as $Q_u(k)$. Thus $Q_u(k) = \mathbb{E}_{y \in \{y^1, \ldots, y^m\}} \hat{q}(F_y; u)$.

We now apply Lemma 1 to obtain Lemma 2, a second useful lemma. It again concerns any Combiner. We let $V_u$ play the role of $\chi$ in Lemma 1, so that $V_u(k)$ plays the role of $t_{\chi}(k)$. Then Lemma 1 implies Lemma 2.

**Lemma 2**

If $u$ is a payoff function which is regular for $I_k$ for all $k \in [0,1]$, then the following statements hold:

1. The value $V_u(k)$ is convex in $k$ on $[0,1]$.
2. Consider the best-quantity function $\hat{q}(\cdot; u)$. If $\hat{q}(\cdot; u)$ is convex (concave) on $\Delta(\Theta)$, then $Q_u(k)$ is convex (concave) in $k$ on $[0,1]$.

We now have the following Theorem. It consists of three statements.

**Theorem 1**

1. Suppose the following holds:

$$\left( + \right) \text{ For every } k \in [0,1] \text{ and every payoff function } u \text{ that is regular for the posteriors of } I_k, \text{ we have } V_u(k) \geq V_u(0).$$

Then the Combiner is a Blackwell IG, i.e., for all $u$ we have $V_u(k') \geq V_u(k)$ whenever $1 \geq k' > k \geq 0$.

2. If the initial anchor structure $I_0$ is null, then the Combiner is a Blackwell IG.
If \( u \) is regular for every \( I_k \), then two statements hold: (a) if \( \hat{q}(\cdot;u) \) is convex on \( \Delta(\Theta) \) and \( Q_u \) is minimized at \( k = 0 \), then \( Q_u(k) \) is convex on \([0,1]\) and is nondecreasing in \( k \); (b) if \( \hat{q}(\cdot;u) \) is concave on \( \Delta(\Theta) \) and \( Q_u \) is maximized at \( k = 0 \), then \( Q_u(k) \) is concave on \([0,1]\) and is nonincreasing in \( k \).

Statement (1) follows from Lemma 1 and 2. Statement (2) follows from statement (1) and the fact that if \( I_0 \) is null, then \( V_u \) is minimized at \( k = 0 \). Statement (3) follows from Lemma 1 and assertion (2) in Lemma 2.

Now consider a payoff function \( u \) which is regular for every \( I_k \). Statement (3) says that whether or not \( I_0 \) is null, we have Complements for the effort measure \( k \) if average best quantity is minimized at \( k = 0 \) and best quantity is convex in the posterior, and we have Substitutes if average best quantity is maximized at \( k = 0 \) and best quantity is concave in the posterior. If \( I_0 \) is null, then, according to statement (2), the Combiner must be a Blackwell IG. So if \( I_0 \) is null then a stronger and simpler statement follows:

We have Complements for the effort measure \( k \) if best quantity is convex in the posterior, and we have Substitutes if best quantity is concave in the posterior.

### 5.1.3 Combiners whose initial anchor structure lacks the null property.

There are Combiners whose initial anchor structure \( I_0 \) lacks the null property but who are nevertheless Blackwell IGs. Consider any Blackwell Combiner for whom \( I_0 \) is null and arbitrarily pick from that Combiner’s structures a non-null structure \( I_{k^*} \) with \( 0 < k^* < 1 \). Now let \( I_{k^*} \) play the role of a new initial anchor structure and let all the structures \( I_k \) with \( k > k^* \) remain as they were. Then we have a new Blackwell combiner.

There are also Combiners whose initial anchor structure \( I_0 \) lacks the null property and are not Blackwell IGs. Consider the following example:

**Example 1** Let \( \Lambda^1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \Lambda^0 = \begin{bmatrix} 1/6 & 5/6 \\ 5/6 & 1/6 \end{bmatrix} \). Note that \( \Lambda^1 \) yields perfect information and is more informative than \( \Lambda^0 \), though \( \Lambda^0 \) is not null. But \( \Lambda^{2/5} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \) is null and hence is less informative than both of the anchors. So this Combiner is not a Blackwell IG.

### 5.1.4 The value curve and the average-quantity curve: some examples.

To visualize more clearly the convexity of \( V_u \) on the Combiner’s effort set \([0,1]\), as well as Statement (4) of Theorem 1, we now graph several examples of the value curve and the average-quantity curve. In both of the figures which follow:
• The Producer’s payoff function is \( u = \theta q - \frac{1}{1+a} q^{1+a} \).

• The state set is \( \Theta = \{ \theta_1, \theta_2, \theta_3 \} = \{0.3, 1.1, 1.7\} \). Each state has prior probability \( \frac{1}{3} \).

• The Combiner’s signal set is \( Y = \{ y_1, y_2, y_3 \} \). The Combiner’s anchor structure \( I_0 \) is null and the anchor structure \( I_1 \) is perfect. Thus

\[
\Lambda^0 = \begin{pmatrix}
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3}
\end{pmatrix}, \quad \Lambda^1 = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

Given a signal \( y \), the Producer’s best quantity is \( \left[ E(\theta \mid y) \right]^{1/a} \), which yields the expected payoff \( \frac{1}{1+a} \cdot \left[ E(\theta | y) \right]^{(1+a)/a} \). We consider three values of the parameter \( a \), namely 0.7, 1.0, 1.8. Averaging over the three signals, we obtain the graphs of value and of average quantity shown in Figures 1 and 2. The shapes of these curves follow from Lemma 2 and Theorem 1.

5.1.5 Choosing the best effort \( k \): corner solutions and interior solutions.

In Figure 1 value rises whenever effort increases. That is guaranteed by Lemma 2, (which establishes the convexity of value as a function of the effort \( k \)), since value is higher at \( k = 1 \) than at \( k = 0 \). Suppose now that effort cost is linear: every effort \( k \in [0, 1] \) costs \( Pk \) dollars, where \( P > 0 \). So the “best” effort maximizes \( V_u(k) - V_u(0) - Pk \). The convexity of the value curves, tells us that if \( P \geq V_u(1) - V_u(0) \), then the best effort is \( k = 0 \); but if \( P < V_u(1) - V_u(0) \), then the best effort is \( k = 1 \). If the best effort is going to lie in the interior of \([0, 1] \), then the cost of effort has to rise nonlinearly. Specifically, if the cost of effort \( k = c(k) \), and if \( k^* \), with \( 0 < k^* < 1 \) is the unique maximizer of \( V_u(k) - V_u(0) - c(k) \), then \( c \) rises more slowly than \( V_u \) at \( k < k^* \) and rises faster than \( V_u \) at \( k > k^* \).

5.1.6 The effect of changes in the Combiner’s effort on the strength of the Complements/Substitutes effect.

Examining Figure 2, we see that for \( a = 0.7 \) we have a strong form of Complements: average quantity is strictly increasing in effort. Moreover, average quantity is strictly convex in effort; average quantity rises more rapidly when effort increases. For \( a = 1.8 \) we have a strong form of Substitutes (average quantity is strictly decreasing in effort). Moreover, average quantity is strictly concave in effort; average quantity falls more rapidly when effort increases. For \( a = 1 \) quantity does not change when effort increases.
Figure 1

- $a = 0.7$
- $a = 1.0$
- $a = 1.8$
Figure 2

average quantity

\( a = 0.7 \)
\( a = 1.0 \)
\( a = 1.8 \)

- \( k \) (IG effort)
The strict convexity (strict concavity) of the average-quantity curves in Figure 2 follows whenever (as in the Figure-2 example): (i) the anchor structures are the null structure and the perfect-information structure, (ii) the expected value of best quantity for a given posterior is a strictly convex (strictly concave) function of the posterior mean, and (iii) each state has the same prior probability. There are other classes of priors for which we again have strictness. In any case Theorem 1 assures us that whatever the priors and the payoff function may be, average quantity is never a strictly concave function of effort in the Complements case and is never a strictly convex function of effort in the Substitutes case.

Strict convexity in the Complements case would mean that the “Complements effect” — the change in average quantity for a small increment in effort — is low when informational effort is small and high when informational effort is large. Strict concavity would mean that the Complements effect is high when informational effort is small and low when informational effort is large. For the Substitutes case we interchange “low” and “high” in the preceding statements.

In the Introduction we provided a principal motivation for studying the Complements/Substitutes question: when is an increase in the number of “information workers” accompanied by an increase (decrease) in the number of “non-information” workers? In an extremely simple version of that question, we can suppose that the Producer has to hire a single input, namely “non-information” labor, to produce his chosen product quantity, where one unit of labor is needed for every unit of product. The Producer responds to the Combiner’s signals. The Combiner uses information workers, and the effort $k$ increases when the information workforce grows. In the Complements case we will see a rise in the non-information workforce when we see a rise in the information workforce. If we have strict convexity, then the effect of a small increment in the information workforce on the size of the non-information workforce is stronger when the information workforce is large. Since strict concavity can be ruled out, the effect is never weaker when the information workforce is large. There are counterparts of these statements for the Substitutes case. In summary: if we know enough about the Producer’s prior and payoff function to verify that the average-quantity curve has one of the shapes in Figure 2, then we can make interesting predictions about workforce composition.

5.1.7 The garbling matrix for the Blackwell Combiner.

We can establish Statement (3) of Theorem 1 in a different way. We can show that the Combiner is a Blackwell IG when $I_0$ is null by explicitly exhibiting the relevant garbling matrix and appealing to the Finite Blackwell Theorem. The garbling matrix is of interest in itself since it provides a clearer view of the way information is degraded when $I_{k'}$ is replaced by $I_k$ and

---

9To show this, we modify the proof of Lemma 1 (given in Appendix A).
0 \leq k < k' \leq 1. We exhibit an \( m \)-by-\( m \) row-stochastic matrix \( B_{kk'} \) such that \( \Lambda^k = \Lambda^{k'} \cdot B_{kk'} \). We find that the garbling matrix \( B_{kk'} \) is a linear combination of the \( m \)-by-\( m \) identity matrix and an \( m \)-by-\( m \) matrix in which all rows are the same. The repeated row has nonnegative elements that sum to one. The first matrix has weight \( k \), which falls when \( k' \) rises (i.e., when the structure \( I_{k'} \) “improves”), while the second matrix has weight \( 1 - k \). In the theorem which follows we let the symbol \( e_t \) denote the \( t \)-by-1 matrix in which every entry is one. Then we may write the \( n \)-by-\( m \) likelihood matrix \( \Lambda^0 \) for the null structure \( I_0 \) as \( e_n \cdot r \), where \( r \) is the row vector \((q_1, q_2, \ldots, q_m)\).

**Theorem 2**

Consider \( k, k' \) with \( 0 \leq k < k' \leq 1 \) and the structures \( I_k, I_{k'} \), whose respective \( n \)-by-\( m \) likelihood matrices are \( \Lambda^k = (1 - k) \cdot \Lambda^0 + k \Lambda^1 \) and \( \Lambda^{k'} = (1 - k') \cdot \Lambda^0 + k' \cdot \Lambda^1 \). Suppose that the structure \( I_0 \) has the null property, i.e., \( \Lambda^0 = e_n \cdot r \), where \( r \) is a row vector \((r_1, \ldots, r_m)\) with nonnegative elements summing to one. Then \( I_k \) is a garbling of \( I_{k'} \). Specifically, the \( m \)-by-\( m \) matrix

\[
B_{kk'} = \frac{k}{k'} H_m + \left( 1 - \frac{k}{k'} \right) e_m \cdot r,
\]

where \( H_m \) denotes the \( m \)-by-\( m \) identity matrix, is row-stochastic and satisfies

\[
\Lambda^k = \Lambda^{k'} \cdot B_{kk'}.
\]

Here is an example illustrating Theorem 2 for three states and two signals.

**Example 2** The two anchors are

\[
\Lambda^1 = \begin{pmatrix} .8 & 0.2 \\ 0.5 & 0.5 \\ 0.2 & 0.8 \end{pmatrix}, \quad \Lambda^0 = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot (0.5 \ 0.5).
\]

Thus \( r \) is the vector \((0.5\ 0.5)\). Now consider \( k = 0.3 \) and \( k' = 0.6 \). Then

\[
\Lambda^{k'} = (1 - k') \Lambda^0 + k' \Lambda^1 = 0.4 \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} + 0.6 \begin{pmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \\ 0.2 & 0.8 \end{pmatrix} = \begin{pmatrix} 0.68 & 0.32 \\ 0.5 & 0.5 \\ 0.32 & 0.68 \end{pmatrix}.
\]

Similarly,

\[
\Lambda^k = (1 - k) \Lambda^0 + k \Lambda^1 = 0.7 \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} + 0.3 \begin{pmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \\ 0.2 & 0.8 \end{pmatrix} = \begin{pmatrix} 0.59 & 0.41 \\ 0.5 & 0.5 \\ 0.41 & 0.59 \end{pmatrix}.
\]

The garbling matrix is

\[
B_{kk'} = \frac{k}{k'} H_2 + \left( 1 - \frac{k}{k'} \right) e_3 \cdot r = 0.5 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 0.5 \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix}.
\]
We have:
\[
\begin{pmatrix}
0.59 & 0.41 \\
0.5 & 0.5 \\
0.41 & 0.59
\end{pmatrix} = \begin{pmatrix}
0.68 & 0.32 \\
0.5 & 0.5 \\
0.32 & 0.68
\end{pmatrix} \cdot \begin{pmatrix}
0.75 & 0.25 \\
0.25 & 0.75
\end{pmatrix},
\]
i.e., \( \Lambda^k = \Lambda^{k'} \cdot B_{kk'} \).

5.1.8 The One-point Forecaster: a special case of the Blackwell Combiner. Suppose now that the number of signals equals the number of states \((m = n)\). Consider an IG whose effort is measured by \(x \in [1/n, 1]\). At effort \(x\), the IG’s \(n\)-by-\(n\) likelihood matrix takes the following simple form:

\[
\hat{\Lambda}^x = \begin{pmatrix}
x & \frac{1-x}{n-1} & \cdots & \frac{1-x}{n-1} & \frac{1-x}{n-1}\\
\frac{1-x}{n-1} & x & \cdots & \frac{1-x}{n-1} & \frac{1-x}{n-1}\\
\cdots & \cdots & \ddots & \cdots & \cdots \\
\frac{1-x}{n-1} & \frac{1-x}{n-1} & \cdots & x
\end{pmatrix}.
\]

We may interpret the signal \(y_j\) as a one-point forecast, namely the forecast that the state will be \(\theta_j\). The probability of a correct forecast is \(x\). Effort is measured by \(x\). At \(x = 1/n\) the likelihood matrix is null, every likelihood is \(1/n\), all posteriors are the same as the prior, and the forecasts are useless. At \(x = 1\), every forecast is perfect. We can interpret the One-point Forecaster as a Combiner whose effort set is \(\{\hat{k}(x) : x \in [1/n, 1]\}\), where \(\hat{k}(x) = (nx - 1)/(n - 1)\). This set is the interval \([0, 1]\), since \(\hat{k}\) is continuous and increasing in \(x\), equals zero when \(x = 1/n\), and equals one when \(x = 1\). At this Combiner’s lowest effort, the likelihood matrix is \(\hat{\Lambda}^0 = \hat{\Lambda}^{\hat{k}(\hat{k})}\). Since that likelihood matrix has the null property, the One-point Forecaster indeed has the Blackwell property. We can apply Theorems 1 and 2. More effort by the One-point Forecaster (weakly) benefits every Producer. If \(1/n \leq x < x' \leq 1\) — which is equivalent to \(0 \leq \hat{k}(x) < \hat{k}(x') \leq 1\) — the information structure defined by \(x'\) is more informative than the structure defined by \(x\). If we use the garbling-matrix formula \(B\) in Theorem 2, we see that the One-point Forecaster’s garbling matrix \(B_{\hat{k}(x'), \hat{k}(x)}\) is row-stochastic and satisfies \(\hat{\Lambda}^{x'} = \hat{\Lambda}^{x} \cdot B_{\hat{k}(x'), \hat{k}(x)}\). Moreover, we have Complements (Substitutes) if the Producer’s best quantity is convex (concave) in the posterior: higher effort means that the average value of the best quantity, over the \(n\) possible forecasts, rises or stays the same (falls or stays the same)\(^{10}\).

5.1.9 The non-finite Combiner and other non-finite Blackwell IGs.

It would be unfortunate if we were unable to study the Complements/Substitutes question for the Blackwell Combiner when the state set is a continuum. It would be equally unfortunate if we

\(^{10}\)The term “all or nothing” has sometimes been used to describe a One-point Forecaster when the signal set is the same as the state set and all states have equal prior probability (or equal prior density). Given a signal \(y\) the true state is \(y\) with probability \(x\). With probability \(1 - x\) all states are equally probable, which the prior already told us. Two papers that consider all-or-nothing structures, but in very different settings, are Johnson and Myatt (2006), and Rajan and Saouma (2006).
could not study the question for the Blackwell IG who partitions such a state set and tells the Producer the set of the partitioning in which the state lies. Infinite state sets arise naturally in many economic models. Once the finiteness of the state set is dropped, serious measure-theoretic difficulties arise in constructing an appropriate version of the three statements that appear in the Finite Blackwell Theorem — “Informativeness”, “Garbling”, and “Convex Functions” — and a proof of their equivalence. In mathematical statistics a substantial comparison-of-experiments literature followed Blackwell’s two papers. Various approaches to dropping the finiteness of the state set were pursued. It is difficult, however, to extract appropriate results from this literature and to assemble them into a unified non-finite theorem. Fortunately we now have a unified theorem, prepared for an economic-theory audience. It is found in a doctoral dissertation by Zhixiang Zhang (2008). Zhang drops the finiteness of the state set and the signal set. He partitions the state set into a finite number of subsets. A sequence of finite partitionings, each a refinement of its predecessor, is considered and the finite theorem is applied to each member of the sequence. A limiting procedure then yields a nonfinite analogue of the finite theorem.

Using the new non-finite Blackwell Theorem we can prove non-finite versions of our two Combiner theorems. Moreover, in Section 7 below we obtain Complements/Substitutes results for a Producer who is a monopolist or a price-taker. The Producer uses a Blackwell IG. In our results the state set may be finite or it may be a continuum. (For example, the set of possible prices facing the price-taker may be finite or it may be an interval). Without the new Blackwell theorem we would remain in the dark about the Complements/Substitutes question for the non-finite case.

5.2 Another Blackwell IG: the Erratic IG.

For this Information Gatherer, the signal set is the union of two sets, called $C$ and $D$. With

[11]A brief survey of the post-Blackwell literature is LeCam (1996). Among the papers mentioned is Strassen (1965), which proves the equivalence of “Garbling” and “Convex Functions” under certain assumptions. Some of the work surveyed by LeCam is found in a lengthy monograph by Torgerson (1991). An earlier version of some of that monograph’s results are found in Torgerson (1976). Other key papers are LeCam (1964), Lehmann (1955), and Lehmann (1988).

[12]In the new theorem “more informative than” now has a weaker definition: for every payoff function and every prior, any expected payoff of the less informative experiment can also be obtained for the more informative experiment by using some action-choosing rule. The garbling matrix, moreover, is now generalized. It is replaced by a Markov kernel from the signal space of the more informative experiment to the signal space of the less informative one.

[13]Amershi (1988) also develops a non-finite Blackwell theorem for a primarily economic audience. In that theorem one finds the equivalence of “Informativeness” and “Garbling” but the theorem is silent on “Convex Functions”. The techniques are quite different than those used by Zhang.

[14]In Marschak, Shanthikummar, and Zhou (2014), we present a version of the Zhang theorem and the measure-theoretic assumptions on which it rests. We then present and prove our non-finite Combiner theorems.
probability $k$ the signal belongs to $C$ and with probability $1 - k$ it belongs to $D$. The IG’s effort is measured by $k$. For example, the IG may use the services of two experts. The first expert provides the signals in $C$ and the second provides the signals in $D$. If the first expert is available, he will be used. If not, the second expert is used and that expert is always available. More effort by the IG means that he induces the first expert, by an appropriate payment, to be available more frequently and thereby increases $k$. This is of particular interest if the first expert is more informative than the second.$^{15}$

We shall consider the finite case. As before, the state set is $\Theta = \{\theta_1, \ldots, \theta_n\}$ and the state probabilities are $g_1, \ldots, g_n$. Let $C$ have $c$ signals and let $D$ have $d$ signals. Let $\bar{\Lambda}^0$ be an $n$-by-$d$ row-stochastic matrix and let $\bar{\Lambda}^1$ be an $n$-by-$c$ row-stochastic matrix.

**Definition 5 (The Erratic IG)** For each $k \in [0, 1]$, the Erratic IG has the information structure $\tilde{I}_k$, which has $c + d$ signals and has the likelihood matrix given by

$$\tilde{\Lambda}^k = \left[ k \cdot \bar{\Lambda}^1 \mid (1 - k) \cdot \bar{\Lambda}^0 \right].$$

Note that $w_j = \sum_{i=1}^n [k \cdot \bar{\Lambda}^1_{ij} \cdot g_i]$ is the marginal probability that the signal is $y_j \in C$ and $w_\ell = \sum_{i=1}^n [(1 - k) \cdot \bar{\Lambda}^0_{\ell i} \cdot g_i]$ is the marginal probability that the signal is $y_\ell \in D$. Let $W_k$ denote the CDF on the signals in $C \cup D$ implied by those marginal probabilities.

Now let $\bar{\pi}^C(y)$ denote the vector of $n$ posterior state probabilities given a signal $y \in C$, and let $\bar{\pi}^D(y)$ denote the vector of $n$ posterior state probabilities given a signal $y \in D$. For $y \in C \cup D$, we define

$$\bar{\pi}(y) \equiv \begin{cases} \bar{\pi}^C(y) & \text{if } y \in C \\ \bar{\pi}^D(y) & \text{if } y \in D. \end{cases}$$

When needed we shall add the subscript $k$, to indicate that we are considering the signals of the structure $\tilde{I}_k$. We then have the symbols $\bar{\pi}^C_k(y), \bar{\pi}^D_k(y), \bar{\pi}_k(y)$. The posterior $F^k_y$ is determined by the vector $\bar{\pi}_k(y)$. Instead of using our usual symbol $E_{F^k_y}$ to denote expectation, it will be useful to use the symbol $E_{\bar{\pi}_k(y)}$. For a regular payoff function $u$, we shall let $\hat{q}(\bar{\pi}(y); u)$ denote the best quantity, i.e., the largest maximizer of $E_{\bar{\pi}_k(y)} u(q, \theta)$. For that payoff function, the value of the structure $\tilde{I}_k = (\{F^k_y\}_{y \in C \cup D}; W_k)$ is $\tilde{V}_u(k) = E_{y \in C \cup D} \left[ E_{\bar{\pi}_k(y)} u \left( \hat{q}(\bar{\pi}(y); u), \theta \right) \right]$. The average best quantity, over all the signals of $\tilde{I}_k$, is $\tilde{Q}_u(k) = E_{y \in C \cup D} \left[ \hat{q}(\bar{\pi}_k(y); u) \right]$. Note that the $(c + d)$-by-$(c + d)$ likelihood matrices for the structures $\tilde{I}_0, \tilde{I}_1$ are, respectively, $\tilde{\Lambda}^0 = [0_{nc} \mid \bar{\Lambda}^0]$ and $\tilde{\Lambda}^1 = [ar{\Lambda}^1 \mid 0_{nd}]$.

---

$^{15}$A model in which decision-makers receive “no additional information” with probability $p$ and useful (but not perfect) information with probability $1 - p$ is studied, in a very different context, by Green and Stokey (2003).
where the symbol $0_{rs}$ denotes the $r$-by-$s$ zero matrix. If $k = 1$, the signals in $D$ do not occur; if $k = 0$, the signals in $C$ do not occur. Thus

$$
\tilde{V}_u(1) = E_{y \in C} \left[ E_{\tilde{\pi}_1^C(y)} u\left(\hat{q}(\tilde{\pi}_1^C(y); u), \theta\right)\right], \quad \tilde{V}_u(0) = E_{y \in D} \left[ E_{\tilde{\pi}_0^D(y)} u\left(\hat{q}(\tilde{\pi}_0^D(y); u), \theta\right)\right].
$$

We now obtain a somewhat degenerate analogue of our Combiner Lemma 1. Consider any function $\chi : \Delta(\Theta) \to IR$. Let the symbol $t_\chi(k)$, used in Lemma 1, now denote the average of $\chi(\tilde{\pi}_k(y))$, over all the signals $y$ in the structure $\tilde{I}_k$.

**Lemma 3**

For any function $\chi : \Delta(\Theta) \to IR$, the function $t_\chi : [0, 1] \to IR$ is linear in $k$.

**Proof:** See Appendix B.

Using Lemma 3, we readily obtain the following theorem.

**Theorem 3**

(1) Suppose $\tilde{I}_1$ is more informative than $\tilde{I}_0$, i.e., for any payoff function $u$ that is regular for both structures we have $\tilde{V}_u(1) \geq \tilde{V}_u(0)$. Then the Erratic IG is a Blackwell IG, i.e., for every payoff function $u$ that is regular for all $\tilde{I}_k$, we have $\tilde{V}_u(k') \geq \tilde{V}_u(k)$ whenever $1 \geq k' > k \geq 0$ (the effort measure $k$ has the informativeness property). Moreover, $\tilde{V}_u(k)$ is linear and nondecreasing in $k \in [0, 1]$.

(2) Consider a payoff function $u$ that is regular for every $\tilde{I}_k$ with $k \in [0, 1]$. If $\tilde{Q}_u(1) \geq \tilde{Q}_u(0)$, then $Q_u(k)$ is linear and nondecreasing on $[0, 1]$. If $\tilde{Q}_u(1) \leq \tilde{Q}_u(0)$, then $Q_u(k)$ is linear and nonincreasing on $[0, 1]$.

Statement (2) yields a Complements/Substitutes result. This time we do not appeal to the “Convex Functions” statement in the Finite Blackwell theorem. If average best quantity is not less at $k = 1$ than at $k = 0$, we have Complements. If average best quantity is not more at $k = 1$ than at $k = 0$, we have Substitutes. Moreover, we again have the “corner” phenomenon that we had for the Combiner: if the Producer pays $Pk$ for the IG effort $k$, where $P > 0$ is fixed, then the Producer chooses either $k = 0$ or $k = 1$. If cost is an appropriate nonlinear increasing function $c$, then best effort is interior. Note that since $\tilde{Q}_u$ is linear, we have a contrast to the Combiner: the strength of the Complements (Substitutes) effect does not change when IG effort grows. Finally, note that the Erratic-IG counterpart of our Combiner value graphs in Figure 1 consists of nondecreasing straight lines. For the Combiner the value graphs are nondecreasing if
we assume that $V_u$ is minimized at $k = 0$ (assumption (+) in Theorem 1). But for the Erratic IG no such assumption is needed, since the graphs are linear and $\tilde{V}_u(0) \leq \tilde{V}_u(1)$.

For a counterpart of Theorem 2 (which concerned the Combiner’s garbling matrix), note that the Erratic IG’s garbling-matrix statement for $1 > k' > k \geq 0$ is

$$\begin{bmatrix}
  k \cdot \tilde{\Lambda}^1 & (1-k) \cdot \tilde{\Lambda}^0
\end{bmatrix} = \begin{bmatrix}
  k' \cdot \tilde{\Lambda}^1 & (1-k') \cdot \tilde{\Lambda}^0
\end{bmatrix} \cdot \begin{bmatrix}
  k \cdot \text{Id}_m & (1-k') \cdot B \\
  0 & \text{Id}_m
\end{bmatrix},$$

where $\text{Id}_m$ is the $m$-by-$m$ identity matrix and $B$ is the garbling matrix that relates $\tilde{\Lambda}^0$ to $\tilde{\Lambda}^1$, i.e., $\tilde{\Lambda}^0 = \tilde{\Lambda}^1 \cdot B$.

5.3: Connections between the Erratic IG and the Combiner IG

There are two interesting connections between the Erratic IG and the Combiner.

- To see the first connection between the two IGs, we observe that $\tilde{\Lambda}^k$ is a convex combination of two anchors:

$$\tilde{\Lambda}^k = \begin{bmatrix}
  k \cdot \tilde{\Lambda}^1 & (1-k) \cdot \tilde{\Lambda}^0
\end{bmatrix} = k \begin{bmatrix}
  \Lambda^1 & 0_{nc}
\end{bmatrix} + (1-k) \begin{bmatrix}
  \Lambda^0 & 0_{nd}
\end{bmatrix},$$

where $0_{nc}, 0_{nd}$ are matrices of zeros. So the Erratic IG is a special case of the Combiner.

Mathematically, we can get the anchor $\tilde{\Lambda}^1 = \begin{bmatrix}
  \Lambda^1 & 0_{nc}
\end{bmatrix}$ by adding columns of zeros in $\Lambda^1$. Then these $\Lambda^1$ and $\tilde{\Lambda}^1$ are informationally equivalent in the Blackwell sense. Similarly $\Lambda^0$ and $\tilde{\Lambda}^0$ are equivalent as well. Notice that the argument here does not require that $c = d$ or that $\tilde{\Lambda}^1$ be more informative than $\tilde{\Lambda}^0$.

- There is a second interesting connection when $\Lambda^1$ and $\Lambda^0$ have the same number of signals, i.e., $c = d = m$. By Lemma 2, the value $V_u$ is convex in $k$ and by Lemma 3, the value $\tilde{V}_u$ is linear in $k$. Moreover,

$$\tilde{V}_u(k) = k\tilde{V}_u(1) + (1-k) \cdot \tilde{V}_u(0) = kV_u(1) + (1-k) \cdot V_u(0) \geq V_u(k)$$

for any payoff function $u$ that is regular for all $\tilde{I}_k$ and all $I_k$. (The final inequality follows from Jensen’s inequality and the convexity of $V_u(\cdot)$). Thus, in the terminology of the Blackwell Theorem, the Erratic IG is more informative than the Combiner. Another way of seeing this is to look at the associated garbling-matrix equality:

$$\Lambda^k = \tilde{\Lambda}^k \cdot \begin{bmatrix}
  \text{Id}_m
\end{bmatrix}.$$
Suppose that the anchors $\bar{\Lambda}^1, \bar{\Lambda}^0$ describe experiments of two experts. Then the garbling-matrix equality tells us, informally, that the Erratic IG’s $\tilde{\Lambda}^k$ contains more information than the Combiner IG’s $\Lambda^k$, since in the Erratic-IG case the identity of the expert is revealed to the Producer, while in the Combiner-IG case, that piece of information is missing.

If we further assume that $\bar{\Lambda}^1$ is more informative than $\bar{\Lambda}^0$, i.e., $\bar{\Lambda}^0 = \bar{\Lambda}^1 \cdot B$ for some $B$, then the Erratic IG is a Blackwell IG. We have the following diagram for any $1 \geq k' > k \geq 0$:

$$
\begin{array}{c}
\tilde{\Lambda}^k' \\
\uparrow \\
\Lambda^{k'} \\
\uparrow \\
\Lambda^k
\end{array}
\begin{array}{c}
\tilde{\Lambda}^k \\
\downarrow \\
\Lambda^k
\end{array}
$$

Here each arrow points to an experiment which is more informative in the Blackwell sense. The vertical arrows are implied by the second connection just described, and the horizontal one is implied by part (1) of Theorem 3. Note that the Combiner IG is not necessarily a Blackwell IG even if we assume $\bar{\Lambda}^0 = \bar{\Lambda}^1 \cdot B$. (See Example 1). Hence there is no arrow from $\Lambda^k$ to $\Lambda^{k'}$.

If, however, the anchor $\bar{\Lambda}^0$ is null, then the Combiner is indeed a Blackwell IG, by part (2) of Theorem 1. In view of part (2) of Theorem 1, we then have the following diagram where $1 \geq k' > k \geq 0$:

$$
\begin{array}{c}
\bar{\Lambda}^1 \\
\uparrow \\
\Lambda^1
\end{array}
\begin{array}{c}
\ldots \\
\downarrow \\
\Lambda^k
\end{array}
\begin{array}{c}
\bar{\Lambda}^k' \\
\uparrow \\
\Lambda^{k'}
\end{array}
\begin{array}{c}
\ldots \\
\downarrow \\
\Lambda^k
\end{array}
\begin{array}{c}
\bar{\Lambda}^{k} \\
\downarrow \\
\Lambda^k
\end{array}
\begin{array}{c}
\ldots \\
\downarrow \\
\bar{\Lambda}^{0}
\end{array}

6. Two non-Blackwell Information-gatherers.

For non-Blackwell IGs we have no general tools comparable to the “Convex Functions” statement in the Blackwell theorems. To obtain Complements/Substitutes results we specify a class of payoff functions, a class of priors, or both. We now consider two non-Blackwell IGs.

6.1 The Scale/location Transformer. Here, in addition to the prior on the state set, there is also a “base” distribution. Each of the IG’s signals, together with the prior, defines a posterior and that posterior is always a scale/location transform of the fixed base. For the signal $y$, let $a(y)$ denote the location parameter and let $b(y)$ denote the scale parameter. The first of two structures requires more IG effort than the second if the average of $b(y)$, over all the signals $y$ of the first structure, is less than the average of $b(y)$ over all the signals of the second structure.\(^{18}\) The IG’s

\(^{18}\) For example, let the state set be $[0, 1]$ and let the prior be uniform on $[0, 1]$. Let the IG’s base distribution be uniform on $[-\frac{1}{2}, \frac{1}{2}]$. The signal $y$ identifies the location coefficient $a(y)$ and the scale coefficient $b(y)$. That pair
set \( \{a(y), b(y) : y \in Y\} \) can be chosen so that the IG becomes a Blackwell Partitioning Refiner. Otherwise the Transformer is a non-Blackwell IG. In either case, the Transformer’s structures articulate well with the payoff function of a certain type of Producer, namely the classic price-taking “news vendor”. The state \( \theta \) is the unknown demand for a product whose price is known to be one. The news vendor has to place an order \( q \) before demand is known. The cost is \( c \) per unit, where \( 0 < c < 1 \). Thus the Producer has the payoff function \( u(q, \theta) = \min(q, \theta) - cq \) and responds to a signal \( y \) by choosing the order \( \hat{q}(y) \) which maximizes \( E(\min(q, \theta)) - cq \) under the posterior defined by \( a(y), b(y) \) and the prior. That maximizer is neither a convex function nor a concave function of the posterior. We cannot use the Blackwell theorem. A specialized argument\(^{19}\) shows, however, that for a wide class of priors there is a critical value of \( c \): if \( c \) is less than that value, then more Transformer effort induces a smaller average order (Substitutes), and if \( c \) exceeds the value, then more effort induces a larger average order (Complements).

6.2 The Equal-probability Partitioner. This IG partitions the state set into \( n \) subsets having equal probability (for a given prior). Higher effort means more sets. Except for the special case where higher \( n \) always means a refinement, the Equal-probability Partitioner is not, in general, a Blackwell IG. It may happen that value drops when effort increases. In Marschak, Shanthikumar, and Zhou (2014) we prove two theorems about the Equal-probability partitioner. They concern Producer payoff functions \( u \) which have the following property: for a given signal \( y \) the best quantity given \( y \) equals the posterior mean raised to some power \( m \), while the value \( V_u(F_y) \) equals the posterior mean raised to some power \( \ell \). In one theorem the state set is \([0, 1]\) and the prior CDF has the form \( G(\theta) = \theta^{\frac{1}{2}} \). (The uniform distribution is an example). In the other theorem the state set is \([k, \infty)\), where \( k > 0 \), and the prior CDF has the form \( G(\theta) = 1 - \left( \frac{k}{\theta} \right)^\delta \), where \( \delta > 0 \). (The Pareto-Levy distribution is an example). Both theorems find that average best quantity (over all signals) strictly decreases if \( m > 1 \) or \( m < 0 \), strictly increases if \( 0 < m < 1 \), and stays the same if \( m = 1 \) or \( m = 0 \). So we have (strict) Complements if \( 0 < m < 1 \) and (strict) Substitutes if \( m > 0 \) or \( m < 0 \). Similarly value (averaged over all signals) strictly decreases if \( \ell > 1 \) or \( \ell < 0 \), strictly increases if \( 0 < \ell < 1 \), and stays the same if \( \ell = 1 \) or \( \ell = 0 \). It turns out — surprisingly — that for a wide class of payoff functions we get exactly the same Complements/Substitutes results and value results as we get if the IG is a Blackwell IG. In particular, if the Producer is a price-taker, cost for the quantity \( q \) is \( \frac{1}{1+k} \cdot q^{1+k} \), and price is \( \theta \in [0, 1] \), then both theorems imply that value yields a transform of the base distribution, namely the uniform distribution on \([a(y) - b(y)/2, a(y) + b(y)/2] \subset [0, 1]\). That becomes the Producer’s posterior. Suppose the IG has a finite signal set \( Y \). The marginal probability of each signal \( y \in Y \) is the prior probability that the state lies in \([a(y) - b(y)/2, a(y) + b(y)/2] \). That probability equals \( b(y) \). So the average of the scale parameters, over all the signals, is \( \sum_{y \in Y} b(y)^2 \). If that average drops when the IG switches from one structure to another, then the new structure requires more effort.

\(^{19}\)The argument is presented in Marschak, Shanthikumar, and Zhou (2015)
strictly rises when the IG works harder, we have Complements if \( k < 1 \), and we have Substitutes if \( k > 1 \). If, on the other hand, price is one but cost is \( \frac{\theta}{1+q} \cdot q^{1+k} \), then both theorems imply that value strictly rises when the IG works harder and that we have Complements. In the Economic Applications section which now follows, we shall see that the same results hold for any prior and any Blackwell IG.

7. Economic applications: some Complements/Substitutes results for specific Blackwell-IG/Producer pairs.

7.1 A Blackwell IG and a Producer whose best quantity depends on the posterior mean.

In these applications the state and signal sets may be finite or nonfinite. In the finite case we appeal to the Finite Blackwell Theorem and in the nonfinite case we appeal to the new Nonfinite Blackwell Theorem. Two of our applications (Theorems 4 and 6) concern a price-taker and the third (Theorem 5) concerns a monopolist. All three proofs (in Appendix B) use a statement about a Producer who uses a Blackwell IG and has a payoff function of the form \( u(q, \theta) = \theta \cdot L(q) + M(q) \). The function \( u \) is supermodular, but that fact alone does not imply Complements/Substitutes. Rather our most general statement is as follows:

Let \( L \) and \( M \) be thrice differentiable. Whether we have Complements or Substitutes depends on the sign of \( L' \) and the sign of the second derivative of \( -\frac{M'}{L'} \). We have Complements if (i) \( L' > 0 \) and \( -\frac{M'}{L'} \) is concave or (ii) \( L' < 0 \) and \( -\frac{M'}{L'} \) is convex. We have Substitutes if (iii) \( L' > 0 \) and \( -\frac{M'}{L'} \) is convex or (iv) \( L' < 0 \) and \( -\frac{M'}{L'} \) is concave.

7.1.1 A price-taking Producer who is uncertain about price. Consider a Producer whose payoff function is \( u(q, \theta) = \theta q - C(q) \). The random state variable \( \theta \) is product price and \( C(q) \) is total cost. The state set is \( \Theta \subseteq \mathbb{R}^+ \). The Producer can choose any nonnegative quantity \( q \). If we make standard assumptions about \( C \), we find that for a signal \( y \), defining a posterior distribution \( F_y \) on \( \Theta \), the Producer’s best quantity is a function of the mean \( E_{F_y} \theta \). The Producer uses a Blackwell IG. In the following Theorem 4 we find, somewhat surprisingly, that if \( C \) is thrice differentiable, then it is the sign of the third derivative that determines whether we have Complements or Substitutes.

Theorem 4

Consider a state set \( \Theta \subseteq \mathbb{R}^+ \) and a Blackwell IG, i.e., an IG with an effort measure that has the informativeness property. Consider a Producer who uses the Blackwell IG, can choose any nonnegative \( q \) and has the payoff function \( u(q, \theta) = \theta q - C(q) \). Suppose that \( C \) is thrice differentiable and at every \( q \geq 0 \), we have \( C'' > 0, C''' > 0 \). Then for that effort measure and any prior on \( \Theta \), the following hold:
(i) If $C''' < 0$ at every $q \geq 0$, we have Complements.

(ii) If $C''' > 0$ at every $q \geq 0$, and $C''(0) = 0$, we have Substitutes.

A class of cost functions meeting the conditions of Theorem 5 is $C(q) = rq^{1+k}/(1 + k), r > 0, k > 0$. Then $C'''(q) = rk(k-1)q^{k-2}$. We have Substitutes if $k > 1$ and we have Complements if $0 < k < 1$.

7.1.2 A monopolist who is uncertain about demand. The next Producer we consider is a monopolist who uses a Blackwell IG, can choose any nonegative $q$, and has the payoff function $u(q, \theta) = qD(q, \theta) - C(q)$. The state set is $\Theta \subseteq R^+$. We consider three scenarios.

First scenario: the demand curve rotates, with the quantity intercept fixed. Specifically, we have:
price $= D(q, \theta) = 1 - \theta q$.

Second scenario: the demand curve rotates, with the price intercept fixed. Specifically, we have:
price $= D(q, \theta) = \theta(1 - q)$.

Third scenario: the demand curve maintains its slope but it shifts up and down. Specifically we have: price $= D(q, \theta) = \theta - q$.

Theorem 5

Consider a state set $\Theta \subseteq R^+$ and a Blackwell IG, i.e., an IG with an effort measure that has the informativeness property. Consider a Producer who uses the Blackwell IG, can choose any nonnegative $q$, and has the payoff function specified in one of the three scenarios. Suppose that the cost function is thrice differentiable. Then for that effort measure and any prior on $\Theta$, the following hold:

First scenario: If $C''' < 0$ at all $q > 0$ we have Complements.

Second scenario: If $C$ satisfies the additional condition $C''(0) = 0$ and if $C''' > 0$ at all $q > 0$ we have Complements.

Third scenario: If $C''' < 0$ at all $q > 0$ we have Complements. If $C''' > 0$ at all $q > 0$ we have Substitutes.

For an example of the three scenarios consider again the cost function $C(q) = rq^{1+k}/(1 + k), k \geq 0, r > 0$. In the first scenario we have Complements if $k \in (0, 1)$. In the second scenario, we have Substitutes if $k > 1$. In the third scenario we have Substitutes if $k > 1$ and Complements if $k < 1$. 

25
7.1.3 A price-taker who is uncertain about cost. Now consider a Producer who sells at a fixed price, which we take to be one, and has total cost function $\theta C(q)$. Payoff is $u(q, \theta) = q - \theta C(q)$ and any nonnegative $q$ can be chosen.

**Theorem 6**

Consider a state set $\Theta \subseteq \mathbb{R}^+$ and a Blackwell IG, i.e., an IG with an effort measure that has the informativeness property. Consider a Producer who uses the Blackwell IG, can choose any nonnegative $q$, and has the payoff function $u(q, \theta) = q - \theta C(q)$. Suppose that the cost function is twice differentiable and satisfies $C(0) = 0$, $C'(0) = 0$, and $C' > 0, C'' > 0$ for all $q \geq 0$. Then for that effort measure we have Complements if $\frac{1}{C'(q)}$ is convex.

For an example consider again $C(q) = r\frac{1}{1+k} q^{1+k}, r > 0, k > 0$. Then $\frac{1}{C'(q)} = \frac{1}{rq^k}$ is convex for $q > 0$. So we have Complements for all $k > 0$. Note that there is no function $C : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, with $C' > 0$, such that $\frac{1}{C'}$ is concave. So if every positive number is a possible posterior mean, then we cannot have Substitutes.

7. Concluding remarks.

We have addressed a well-motivated question that appears to have escaped intensive study: when does more informational effort lead to higher quantity produced? The question can be posed in a large and varied collection of real settings. We have considered several classes of IGs and Producers. The closest we get to generality are propositions about Blackwell IGs. But there are interesting IG/Producer pairs for whom an IG effort measure with the informativeness property does not exist, and there are interesting pairs for which it does exist, but the Producer’s best quantity is not a convex or concave function of the posterior. In our economic applications, we obtain sharp Complements/Substitutes results when best quantity depends on the posterior mean and the Producer uses a Blackwell IG. The critical role of the sign of the third derivative of the cost function, when the Producer is a monopolist or price-taker, is an unexpected result, not suggested by any simple intuition. For the non-Blackwell Equal-probability Partitioner, and certain priors and Producer payoff functions, we get Complements/Substitutes results that duplicate what we find for those payoff functions when the IG is a Blackwell IG.

---

20Suppose $\frac{1}{C'}$ were a concave function, say $f$, on $\mathbb{R}^+$. Then $f$ is bounded from above by any tangent line. The tangent line is negatively sloped and the absolute value of its slope becomes arbitrarily large as $t$ increases without limit. Hence we have $\lim_{t \to \infty} f(t) = -\infty$. That contradicts our assumption that $C' > 0$.

21In Marschak, Shanthikumar, and Zhou (2014) we present an example where best quantity depends on the posterior variance as well as the posterior mean. The Producer constructs a portfolio and chooses the amount to spend on a risky asset; the state $\theta$ is its yield. He maximizes a quadratic utility-of-wealth function. We have Complements if he uses the One-point Forecaster or the Equal-probability Partitioner: when these IGs work harder he spends more on the risky asset.
A more thorough study would include incentive issues. The IG would be free to choose effort (which may be hidden) as well as the signal he reports, and would be rewarded by the Producer. The Producer seeks a reward scheme that is best for him among all the truth-inducing schemes to which the IG will agree. We would then have a new Complements/Substitutes question: if the IG’s cost for every effort drops (because technology improves or information workers’ wages drop), will there be a rise or a fall in the Producer’s average quantity when he responds to the signals that the IG sends him under a best reward scheme? One can certainly worry that the answer to such a difficult question might reverse the direction of a Complements/Substitutes result that we obtain in our much simpler incentive-free framework.

We conclude by emphasizing again the strong empirical motivation for building a Complements/Substitutes theory. To explain and predict the occupational shifts and the productivity effects of the ongoing IT revolution requires some clarifying theory. One approach is the study of skills and tasks that has been pursued by labor economists. Our IG/Producer model is another. Our results suggest that a useful though difficult path for empiricists is the detailed observation of individual real IGs and the tasks they perform, the measurement of their efforts, and the actions which real Producers take in response to the signals received from the IG that they choose. One fruit of theory is the identification of datasets that would be useful if they were assembled. Without theory they might be overlooked.
Appendix

A: Proofs of the Combiner results.
B: Proof of Lemma 3 (concerning the erratic IG).
C: Proofs of the theorems about economic applications where the IG is a Blackwell IG.

Appendix A: Proofs of the Combiner results.

Proof of Lemma 1

Suppose \( \chi \) is convex. The expected value \( \eta_\chi(k) \) is a sum of \( m \) terms, one for each signal. In this proof we will not need the subscript \( j \) which identifies a particular signal or the symbol \( u_j^k \), which denotes the probability of the \( j \)th signal in the structure \( I_k \). Instead we observe that each of the \( m \) terms has the form \( w(k) \cdot \chi(\frac{\tilde{u}(k)}{w(k)}) \). Here \( \tilde{u} \) is a vector with \( n \) components, one for each state; \( w(k) > 0 \); and both \( \tilde{u} \) and \( w \) are linear in \( k \). It suffices to prove that \( w(k) \cdot \chi(\frac{\tilde{u}(k)}{w(k)}) \) is convex in \( k \). Consider \( k_1, k_2 \in [0, 1], \alpha \in [0, 1], \) and \( k = \alpha k_1 + (1 - \alpha) \cdot k_2 \). We have to show that

\[
\chi \left( \frac{\tilde{u}(k)}{w(k)} \right) \cdot w(k) \leq \alpha w(k_1) \cdot \chi \left( \frac{\tilde{u}(k_1)}{w(k_1)} \right) + (1 - \alpha) \cdot w(k_2) \cdot \chi \left( \frac{\tilde{u}(k_2)}{w(k_2)} \right)
\]

(1)

The linearity of \( \tilde{u} \) and \( w \) imply that \( \tilde{u}(k) = \alpha \tilde{u}(k_1) + (1 - \alpha) \cdot \tilde{u}(k_2) \), \( w(k) = \alpha w(k_1) + (1 - \alpha) \cdot w(k_2) \). Hence

\[
\chi \left( \frac{\tilde{u}(k)}{w(k)} \right) \cdot w(k) = \chi \left( \frac{\alpha \tilde{u}(k_1) + (1 - \alpha) \cdot \tilde{u}(k_2)}{\alpha w(k_1) + (1 - \alpha) \cdot w(k_2)} \right) \cdot (\alpha w(k_1) + (1 - \alpha) \cdot w(k_2))
\]

\[
= \chi \left( \frac{\alpha w(k_1)}{\alpha w(k_1) + (1 - \alpha) \cdot w(k_2)} \cdot \frac{\tilde{u}(k_1)}{w(k_1)} + \frac{(1 - \alpha) \cdot w(k_2)}{\alpha w(k_1) + (1 - \alpha) \cdot w(k_2)} \cdot \frac{\tilde{u}(k_2)}{w(k_2)} \right)
\]

\[
\cdot (\alpha w(k_1) + (1 - \alpha) \cdot w(k_2))
\]

\[
\leq \left[ \frac{\alpha w(k_1)}{\alpha w(k_1) + (1 - \alpha) \cdot w(k_2)} \cdot \chi \left( \frac{\tilde{u}(k_1)}{w(k_1)} \right) + \frac{(1 - \alpha) \cdot w(k_2)}{\alpha w(k_1) + (1 - \alpha) \cdot w(k_2)} \cdot \chi \left( \frac{\tilde{u}(k_2)}{w(k_2)} \right) \right]
\]

\[
\cdot (\alpha w(k_1) + (1 - \alpha) \cdot w(k_2))
\]

\[
= \alpha w(k_1) \cdot \chi \left( \frac{\tilde{u}(k_1)}{w(k_1)} \right) + (1 - \alpha) \cdot w(k_2) \cdot \chi \left( \frac{\tilde{u}(k_2)}{w(k_2)} \right).
\]

The inequality which comes after the first two equalities follows from (i) the convexity of \( \chi \), (ii) the fact that \( w(k) > 0 \) (which implies that \( \alpha w(k_1) + (1 - \alpha) \cdot w(k_2) > 0 \)), and (iii) the fact that \( \frac{\alpha w(k_1)}{\alpha w(k_1) + (1 - \alpha) \cdot w(k_2)} \) and \( \frac{(1 - \alpha) \cdot w(k_2)}{\alpha w(k_1) + (1 - \alpha) \cdot w(k_2)} \) are nonnegative and sum to one. The proof when \( \chi \) is concave is analogous.

We now turn to the second statement in the Lemma. Consider the case where \( t_\chi \) is convex and \( t_\chi \) is minimized at \( k = 0 \). For \( k_1 > 0 \), we have \( t_\chi(k_1) \geq t_\chi(0) \). If \( k_1 > k_2 > 0 \), then the convexity
of $t_\chi$ implies that \( \frac{t_\chi(k_1) - t_\chi(k_2)}{k_1 - k_2} \geq \frac{t_\chi(k_1) - t_\chi(0)}{k_1 - 0} \). Since the second fraction is nonnegative and \( k_1 - k_2 > 0 \), we obtain $t_\chi(k_1) \geq t_\chi(k_2)$. There is an analogous proof for the case where $t_\chi$ is concave and $t_\chi$ is maximized at $k = 0$. \(\square\)

**Proof of Theorem 3.** We first show that $B_{kk'}$ is row-stochastic, i.e., all its entries are non-negative and $B_{kk'} \cdot e_m = e_m$. Clearly the entries are nonnegative. Since $q \cdot e_m = 1$, we have:

$$B_{kk'} \cdot e_m = \left[ \frac{k}{k'} H_m + \left( 1 - \frac{k}{k'} \right) e_m \cdot q \right] \cdot e_m = \frac{k}{k'} H_m \cdot e_m + \left( 1 - \frac{k}{k'} \right) \cdot e_m \cdot (q \cdot e_m) = \frac{k}{k'} e_m + \left( 1 - \frac{k}{k'} \right) \cdot e_m = e_m$$

Next note that since every row of a likelihood matrix sums to one, we have,

$$\sum_{i=1}^{n} H_{ij} e_m = \sum_{i=1}^{n} \lambda_{ij} e_m = \lambda^0 \cdot e_m. \text{Using that fact, and recalling that } \lambda^0 = e_n \cdot q, \text{we have:}$$

$$\Lambda^{k'} \cdot B_{kk'} = (k' \Lambda^1 + (1 - k') \cdot \Lambda^0) \cdot \left[ \frac{k}{k'} H_m + \left( 1 - \frac{k}{k'} \right) e_m \cdot q \right] = k' \Lambda^1 \cdot \frac{k}{k'} H_m + (1 - k') \cdot \Lambda^0 \cdot \left( 1 - \frac{k}{k'} \right) \cdot e_m \cdot q + (1 - k') \cdot \Lambda^0 \cdot \left( 1 - \frac{k}{k'} \right) \cdot e_m \cdot q = k \Lambda^1 + (1 - k') \cdot \frac{k}{k'} \cdot \Lambda^0 \cdot e_n \cdot q = k \Lambda^1 + \left[ (1 - k') \cdot \frac{k}{k'} + k' \cdot \left( 1 - \frac{k}{k'} \right) + (1 - k') \cdot \left( 1 - \frac{k}{k'} \right) \right] \cdot \Lambda^0 = k \Lambda^1 + (1 - k) \cdot \Lambda^0 = \Lambda^k$$

That concludes the proof. \(\square\)

**Appendix B: proof of Lemma 3 (concerning the Erratic IG).**

For the structure $I_k$, consider the signal $y_j \in C$. Its probability is $\sum_{i=1}^{n} [k \cdot \bar{\lambda}_{ij} \cdot p_i]$. Its likelihood given the state $\theta_i$ is $k \cdot \bar{\lambda}_{ij}$. Given the signal $y_j$, the posterior probability of the state $\theta_i$ is

$$\frac{p_i \cdot k \cdot \bar{\lambda}_{ij}^1}{\sum_{i=1}^{n} k \cdot p_i \cdot \bar{\lambda}_{ij}} = \frac{p_i \bar{\lambda}_{ij}^1}{\sum_{i=1}^{n} p_i \cdot \bar{\lambda}_{ij}}.$$ 

So the vector of $n$ state posteriors given the signal $y_j \in C$ is $\bar{\pi}^C(y_j) \in \Delta$, where

$$\bar{\pi}^C(y_j) = \frac{1}{\sum_{i=1}^{n} p_i \cdot \bar{\lambda}_{ij}^1} (p_1 \cdot \bar{\lambda}_{ij}^1, \ldots, p_n \cdot \bar{\lambda}_{nj}^1).$$

Similarly, consider the signal $y_i \in D$. Its probability is $\sum_{i=1}^{n} [(1 - k) \cdot \bar{\lambda}_{i0}^0 \cdot p_i]$. Its likelihood given the state $\theta_i$ is $(1 - k) \cdot \bar{\lambda}_{i0}^0$. The posterior probability of the state $\theta_i$ given the signal $y_i \in D$ is

$$\sum_{i=1}^{n} (1 - k) \cdot p_i \cdot \bar{\lambda}_{i0}^0 = \frac{p_i \bar{\lambda}_{i0}^0}{\sum_{i=1}^{n} p_i \cdot \bar{\lambda}_{i0}^0}.$$

29
The vector of $n$ state posteriors given $y_\ell \in D$ is $\pi^D(y_\ell) \in \Delta$, where

$$\pi^D(y_\ell) = \frac{1}{\sum_{i=1}^{n} p_i \cdot \Lambda^0_{i\ell}} (p_1 \cdot \Lambda^0_{1\ell}, \ldots, p_n \cdot \Lambda^0_{n\ell}).$$

The expected value of $\chi$, over all the signals in the structure $I_k$, is

$$\eta_{\chi}(k) = \sum_{y_j \in C} \text{prob. of } y_j \cdot \chi(\pi_C(y_j)) + \sum_{y_\ell \in D} \text{prob. of } y_\ell \cdot \chi(\pi^D(y_\ell))$$

$$= k \cdot \sum_{y_j \in C} \left[ \sum_{i=1}^{n} \Lambda^1_{ij} \cdot p_i \right] \cdot \chi(\pi_C(y_j)) + (1 - k) \cdot \sum_{y_\ell \in D} \left[ \sum_{i=1}^{n} \Lambda^0_{i\ell} \cdot p_i \right] \cdot \chi(\pi^D(y_\ell)).$$

So, as claimed, $\eta_{\chi}$ is linear in $k$ for all functions $\chi$. \qed

Appendix C: proofs of the Economic Applications theorems.

In theorems 4, 5, 6 the Producer can choose any quantity in $\mathbb{R}^+$. We are given a state set $\Theta \subseteq \mathbb{R}$ and a Producer’s payoff function $u : \mathbb{R}^+ \times \Theta \to \mathbb{R}$. The Producer’s best quantity depends only on the mean of the posterior. Consider three functions:

- The function $\hat{q}(\cdot; u) : \Delta(\Theta) \to \mathbb{R}^+$, where, as before, $\hat{q}(F; u)$ denotes the Producer’s best quantity when $\theta$ has the CDF (or measure) $F$.
- A function $q^* : \mathbb{R} \to \mathbb{R}^+$ such that for any $F \in \Delta(\Theta)$ we have $q^*(E_F \theta) = \hat{q}(F; u)$.
- The function $\tilde{q} : \Delta(\Theta) \to \mathbb{R}^+$ which satisfies $\tilde{q}(F) = q^*(E_F \theta)$.

Since the mean of a CDF (or measure) is linearly related to the CDF (or measure) itself, we have:

(§) If $\hat{q}(\cdot; u)$ is convex (concave) on $\Delta(\Theta)$ then so is $\tilde{q}$.

But it is also true that

(§§) If $q^*$ is convex (concave) on $\mathbb{R}^+$, then $\tilde{q}$ is convex (concave) on $\Delta(\Theta)$.

Recall that the Blackwell Theorems assert that the “Convex Functions” statement implies the “Informativeness” statement. So the statements (§) and (§§) imply that if we are given a prior on $\Theta$, two IG structures $I' = (\{F'_y\}_{y \in Y'}, W'_{Y'})$, $I = (\{F_y\}_{y \in Y}, W_Y)$, and an effort measure with the informativeness property, then the following statement holds:

If $I'$ requires higher effort than $I$ and $q^*$ is convex (concave) on $\mathbb{R}$, then $E_{y \in Y'} \hat{q}(F'_y; u)$ is not less than (is not more than) $E_{y \in Y} \hat{q}(F_y; u)$.
Informally: if \( q^* \) is convex (concave) on \( \mathbb{R} \), then we have Complements (Substitutes), i.e., higher IG effort cannot decrease (increase) average best quantity. In each of the three theorems, we show that the assumptions on the cost function \( C \) imply that for every \( F \in \Delta(\Theta) \) there is a unique best quantity \( \hat{q}(F;u) \) and that there is a function \( q^* : \mathbb{R} \rightarrow \mathbb{R}^+ \) for which \( q^*(E_F \theta) = \hat{q}(F;u) \).

To find \( q^* \), we consider the first-order conditions that a best quantity must satisfy. In all three theorems the payoff function has the form \( u(q,\theta) = \theta \cdot L(q) + M(q) \). It will be helpful to write the expected payoff given a signal in the general form \( \pi(q,w) = wL(q) + M(q) \), where \( w \) denotes the posterior mean given the signal. We shall see that in all three theorems the assumptions made about the cost function \( C \) imply that \( L \) and \( M \) are differentiable and that \( \pi(\cdot,w) \) has a unique maximizer \( q^*(w) \), where \( q^* : \mathbb{R} \rightarrow \mathbb{R}^+ \) and for every \( F \in \Delta(\Theta) \) and every signal \( y \), we have \( q^*(E_{F_y} \theta) = \hat{q}(F_y;u) \).

First we show that:

(A1) \( q^* \) is (weakly) increasing if \( L'(q) > 0 \) for all \( q \geq 0 \).

(A2) \( q^* \) is (weakly) decreasing if \( L'(q) < 0 \) for all \( q \geq 0 \).

To establish these statements, note that \( \partial^2 \pi(q,w)/\partial q \partial w = L'(q) \). If \( L'(q) > 0 \) for all \( q \geq 0 \), then \( \pi \) is supermodular and a standard result\(^{22}\) tells us that \( q^* \) is (weakly) increasing. Suppose, on the other hand that \( L'(q) < 0 \) for all \( q \geq 0 \). Define a new variable \( \tau = -w \) and consider \( \partial^2 \pi(q,-\tau)/\partial q \partial \tau = -L'(q) > 0 \). Then we again have supermodularity. The function \( q^*(-\tau) \) is (weakly) increasing in \( \tau \) and hence \( q^*(w) \) is (weakly) decreasing in \( w \).

Having established (A1),(A2), we shall now use the following two facts:

(B1) For an increasing function \( f \), the inverse is convex if \( f \) is concave and the inverse is concave if \( f \) is convex.

(B2) For a decreasing function \( g \), the inverse is convex if \( g \) is convex and the inverse is concave if \( g \) is concave.

Now let \( h \) denote the inverse of \( q^* \). Since \( q^*(w) \) is the unique solution to the first-order condition \( L'(q) \cdot w + M'(q) = 0 \), we have \( w = h(q^*(w)) \), where \( h(q) = -M'(q)/L'(q) \) if \( L'(q) \neq 0 \). Applying (A1), (A2),(B1), (B2), we obtain the following.

C. Suppose \( L'(q) > 0 \) for all \( q \geq 0 \).

Then

\(^{22}\)See Milgrom and Shannon(1994).
The function \( q^* \) and its inverse \( h = -\frac{M'}{L'} \) are increasing.

\( q^* \) is convex if \( h \) is concave and \( q^* \) is concave if \( h \) is convex.

### D. Suppose \( L'(q) < 0 \) for all \( q \geq 0 \)

Then

- \( q^* \) and its inverse \( h = -\frac{M'}{L'} \) are decreasing.
- \( q^* \) is convex if \( h \) is convex and \( q^* \) is concave if \( h \) is concave.

We now apply (B1), (B2), (C) and (D) to the Producers who occur in Theorems A1-A2.

**Proving Theorem 4** This Producer’s payoff function is \( u(q, \theta) = \theta q - C(q) \), where price is \( \theta \in \Theta \subseteq \mathbb{R}^+ \) and any nonnegative quantity \( q \) can be chosen. So in the general framework just discussed, the term \( L(q) \) is \( q \) and \( M(q) = -C(q) \). By assumption, \( C \) is twice differentiable, with \( C(0) = 0 \) and \( C'(q) > 0, C''(q) > 0 \) for all \( q \geq 0 \). Then given the posterior mean \( w \), the Producer chooses \( q^*(w) \), the unique maximizer of \( \pi(q, w) = L(q) \cdot w + M(q) \). Since \( L'(q) = 1 > 0 \) for \( q \geq 0 \) we can apply (C1), (C1). Note that \( h(q) = -\frac{M'(q)}{L'(q)} = \frac{C'(q)}{1} = C'(q) \).

Suppose now that the function \( C \) is thrice differentiable. Then, using (C1), (C2), we conclude that the function \( q^* : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \) is concave if \( C'''(q) > 0 \) for all \( q \geq 0 \) and convex if \( C'''(q) < 0 \) for all \( q \geq 0 \). Hence for this Producer and any Blackwell IG: We have **Substitutes** if \( C'''(q) > 0 \) for all \( q \geq 0 \); we have **Complements** if \( C'''(q) < 0 \) for all \( q \geq 0 \). That is what Theorem 6 asserts.

**Proving Theorem 5** The state variable \( \theta \) is nonnegative. By assumption, the cost function \( C \) is twice differentiable and satisfies \( C(0) = 0, C' > 0, C'' > 0 \) for all \( q \geq 0 \).

**First scenario:**

Given a posterior mean \( w \), the Producer maximizes

\[
\pi(q, w) = (-q^2) \cdot w + (q - C(q)).
\]

So \( L(q) = -q^2 \) and \( M(q) = q - C(q) \). For every \( w > 0 \), there is a unique maximizer \( q^*(w) \). Since \( L' < 0, (D1) \) tells us that \( q^* \) is decreasing. Make the further assumption that \( C'(0) = 0 \). Then the first-order condition for a maximizer of \( \pi(\cdot, w) \) tells us that \( 1 - C'(q^*(w)) > 0 \). For the inverse function \( h \) we have \( h(t) = -\frac{M'(t)}{L'(t)} = \frac{1-C'(t)}{2t} \) and \( h''(t) = \frac{-C''(t)t^2+2C'''(t)+2(1-C'(t))}{2t^3} \). Since \( 1 - C'(q^*(w)) > 0, C'' > 0 \), we conclude that if \( C''' < 0 \), then the inverse of the function \( q^* \) is convex and hence, by (D2), \( q^* \) is also convex.
So for this Producer and any Blackwell IG: if $C'''(q) < 0$ for all $q > 0$, then we have Complements. That is what Theorem 7 asserts for the first scenario.

**Second scenario**

Given the posterior mean $w$, the Producer maximizes $\pi(q, w) = w \cdot (1 - q) \cdot q - C(q)$. We have the additional assumption that $C'(0) = 0$. We permit the Producer to choose any nonnegative $q$. Note that for every $w \geq 0$ there is a unique maximizer of $\pi(q, w)$. As before, the maximizer is denoted $q^*(w)$. The first-order condition is

$$w \cdot (1 - 2q^*(w)) + C'(q^*(w)) = 0$$

Since $C' > 0$, we have $1 - 2q^*(w) > 0$ and for any $w \geq 0$, the maximizer $q^*(w)$ lies in $[0, \frac{1}{2})$. Then $h(t) = \frac{C'(t)}{1 - 2t}, t \in [0, \frac{1}{2})$ is the inverse of the function $q^*$. We have $h'(t) = \frac{(1 - 2t)C'' + 2C'}{(1 - 2t)^2} > 0$, since $1 - 2t > 0, C' > 0, C'' > 0$. We then obtain $h''(t) = \frac{(1 - 2t)^2C''' + 4(1 - 2t)C'' + 8C'}{(1 - 2t)^3}$. If $C''' > 0$, then $h'' > 0$, so $h$ is convex on $[0, \frac{1}{2})$. Therefore (by (B1)), $q^*$ is increasing and concave if $C''' > 0$. We conclude that for any Blackwell IG: we have Substitutes if $C'''(q) > 0$ for all $q \geq 0$. That is what Theorem 7 asserts for the second scenario.

**Third scenario**

Now we have $\pi(q, w) = qw - q^2 - C(q)$. So $L(q) = q, M(q) = -q^2 - C(q)$. We have $L'(q) = 1 > 0, h(q) = -\frac{M(q)}{L'(q)} = 2q + C'(q)$, and

$$h''(q) = C'''(q).$$

Applying (C1), we see that the unique maximizer $q^*(w)$ is increasing in $w$. Since $h$ is convex if $C''' > 0$ and concave if $C''' < 0$, we use (C2) and we conclude that for this Producer and for any Blackwell IG: we have Substitutes if $C'''(q) > 0$ for all $q \geq 0$; we have Complements if $C'''(q) < 0$ for all $q \geq 0$. That is what Theorem 7 asserts for the third scenario.

**Proving Theorem 6.** We again make the additional assumption that $C'(0) = 0$. Given a posterior mean $w$, the producer maximizes $\pi(q, w) = (-C(q)) \cdot w + q$. So $L(q) = -C(q), M(q) = q$, and $h(q) = -\frac{M'}{L'} = \frac{1}{C'(q)}$. For every $w > 0$, we see that $q^*(w)$, the unique maximizer of $\pi$ is positive. Since $L' < 0$ for all $q > 0$, we apply (C2). We conclude that for this Producer and any Blackwell IG we have Complements if $\frac{1}{C'(q)}$ is convex. That is what Theorem 8 asserts.

**REFERENCES**


