A Newsvendor Who Chooses Informational Effort

Thomas Marschak

Walter A. Haas School of Business, University of California, Berkeley, California 94720, USA, marschak@haas.berkeley.edu

J. George Shanthikumar

Krannert School of Management, Purdue University, West Lafayette, Indiana 47907, USA, jshanthi@purdue.edu

Junjie Zhou

School of International Business Administration, Shanghai Uiversity of Finance and Economics, Shanghai 20043, China, zhoujj03001@gmail.com

We study a newsvendor who can acquire the services of a forecaster, or, more generally, an Information Gatherer (IG) to improve his information about demand. When the IG's effort increases, does the average *ex ante* order quantity rise or fall? Do average *ex post* sales rise or fall? Improvements in information technology and in the services offered by forecasters provide motivation for the study of these questions. Much depends on our model of the IG and his efforts. We study an IG who sends a signal to a classic single-period newsvendor. The signal defines the newsvendor's posterior probability distribution on the possible demands and the newsvendor uses that posterior to calculate the optimal order. Each of the possible posteriors is a scale/location transform of the same base distribution. When the IG works harder, the average scale parameter drops. Higher IG effort is always useful to the newsvendor. We show that there is a critical value of order cost. For costs on one side of this value more IG effort leads to a higher average *ex ante* order and for costs on the other side to a lower average order. But for all costs, more IG effort leads to higher average *ex post* sales. We obtain analogous results for a "regret averse" newsvendor who suffers a penalty that is a nonlinear function of the discrepancy between quantity ordered and true demand.

A growing demand-forecasting "industry" (whose current players include Oracle Retail Demand Forecasting, Sylvon, Direct Tech, Demand World) offer individually designed demand forecasts to retailers and inventory managers. Continual improvement in the quality of forecasts is claimed. The technology for assembling datasets on which forecasts can be based (scanner data, survey data) steadily improves. In this paper we study the effects of more and better demand information on the behavior of inventory managers.

Specifically, we consider a model in which the inventory manager is a classic newsvendor whose demand changes at regular intervals and is independent of previous demands. The newsvendor uses his information about forthcoming demand to place an order with a manufacturer. Every unit ordered has the same cost and every unit sold has the same price. The newsvendor's information about forthcoming demand is supplied by an Information-gatherer (IG). After exerting effort, the IG sends one of several possible signals to the newsvendor. The signal determines a posterior demand distribution and the newsvendor uses that distribution to calculate that period's optimal order. Without the IG, the newsvendor would use a prior distribution. When the IG exerts more effort in his study of future demand, the posterior distributions become, on the average, more useful to the newsvendor.

Our central questions are the following: Does more effort on the part of the IG increase or decrease the newsvendor's average *ex ante* order, when we average over all of the IG's possible signals? Does it increase or decrease the newsvendor's average *ex post* sales?

Our paper provides a new approach to a puzzle considered in a number of earlier papers: are inventories and information substitutes? When information about demand improves do inventories shrink? In our approach, inventory size is measured by the newsvendor's average order and better information means higher IG effort. We study a model of the IG in which effort has a natural definition.

Consider a specific example, where the role of "newsvendor" is played by an NGO (non-governmental organization). The NGO orders medicines for a particular disease and ships them to countries where the disease is endemic. Each dose of medicine cures the disease and costs c. Each cured person has value $v > c$. The NGO procures the services of an IG, an expert on the disease, whose effort (research budget) can be varied. The IG provides one of several possible signals, where each signal defines a posterior distribution on forthcoming demand (number of diseased persons). The NGO uses that posterior to choose an expected-cost-minimizing order, where cost is the sum of overage and underage costs. The NGO has to tell its suppliers, who must plan ahead, what its average order will be, and also wants to announce to the world the average number of cured persons. So the NGO wants to know the average, over all the IG's signals, of the *ex ante* expected-cost-minimizing order, as well as the average, over all signals, of the *ex post* deliveries ("sales").

We shall avoid the ambitious question of *how much* IG effort the newsvendor should buy from a member of the forecasting "industry". That would require us to specify a price for each of an IG's possible efforts and to balance the cost of more IG effort against its benefits to the newsvendor. It is natural, in our view, to defer the more ambitious investigation until we have

studied a variety of IG models and we know more, for each of the models, about the direction in which the *ex ante* and *ex post* quantities move when IG effort increases.

There are a number of ways to model an IG and his efforts and we shall have to choose one of them. The IG might partition the possible demands, reporting to the newsvendor the set of the partitioning in which the current demand lies. Higher IG effort might yield a refinement of the partitioning that the IG obtains for a lower effort. Alternatively, the possible partitionings might not be successive refinements. Instead — assuming that the possible demands are the points of an interval — the sets of a partitioning might be subintervals having equal prior probability, where the number of such subintervals grows when IG effort increases. In another model, the IG samples from the population of potential demanders, each of whom, let us say, demands either one unit or two units in the next period. The IG then reports the sampling results to the newsvendor. Higher IG effort means a larger sample. In yet another model, the IG issues a point-valued forecast of the demand and higher effort means that the forecast's average squared error is reduced. Each model of the IG presents its own challenges when we study the effect of more IG effort on the newsvendor's average *ex ante* order quantity and on average *ex post* sales.

We shall identify a special property of the newsvendor's payoff function: it is shift-invariant and satisfies a homogeneity condition. We choose a model of the IG which turns out to articulate well with that special property, so that sharp results can be obtained. In our model, the IG is a *scale/location transformer*. He has a *base distribution* on the possible demands and a collection Y of possible *signals*. For a given value of true demand, there is a probability distribution on Y. After devoting effort to the study of forthcoming demand, the IG obtains a *signal*, say y, and conveys that signal to the newsvendor. The signal y defines a pair $(a(y), b(y))$, where $a(y)$ is a location parameter and $b(y) > 0$ is a scale parameter. The location/scale pair $(a(y), b(y))$ defines a transform of the IG's base distribution and that transform *becomes the newsvendor's posterior*. The newsvendor uses the posterior to calculate the optimal order quantity. The variance of the transformed distribution is increasing in $b(y)$. So it is natural to define the IG's effort in terms of *average-scale reduction* and to say that the IG works harder when the average value of $b(y)$, over all the IG's signals y, drops.

We first consider a classic newsvendor, who sells at a price of one, obtains no salvage value from unsold stock, and has an order cost of c per unit, where $0 < c < 1$. For a given signal y he chooses an order quantity q that maximizes the conditional expectation $E\left(\min(D, q) - cq \mid y\right)$, where D denotes realized demand. We shall obtain results that hold no matter what the IG's base distribution may be and no matter what his collection of possible location/scale pairs may be. Here are our results for the classic newsvendor:

[1] When the IG works harder, the newsvendor benefits. His average profit (weakly) rises .

[2] There is a critical value of the cost c. If cost exceeds the critical value then the newsvendor's average *ex ante* order rises whenever the IG works harder; if cost is less than the critical value then the average order drops whenever the IG works harder. If cost equals the critical value, the average order stays the same whenever the IG works harder.

[3] Whenever the IG works harder, the average of the classic newsvendor's *ex post* sales goes up or stays the same, whatever c may be.

Note that we could, if we wish, avoid the term "works harder", which refers to the effort that the IG chooses. In view of statement **[1]**, we could use instead the term "whenever the newsvendor benefits", which refers to the outcome of the effort choice. But we find it instructive to identify a plausible IG-effort measure (average scale reduction in our location/scale model of the IG) and to interpret our results as telling us the effect of more IG effort on average newsvendor order quantity and average sales. Observe that **[2]** implies

[2′] For any <u>fixed</u> c, the effect of more IG effort on the average *ex ante* order is monotonic: either more effort increases the average order (or leaves it unchanged), or more effort decreases the average order (or leaves it unchanged).

After studying the classic newsvendor, we turn to a "generalized" newsvendor, who suffers a penalty that depends, nonlinearly and asymmetrically, on the gap between the order quantity chosen and the quantity that turns out to be ideal. Such a newsvendor is shown to be "regret averse". For the generalized newsvendor, we obtain counterparts of the preceding results.

What about an IG who is *not* our scale/location transformer? There may be other IGs for whom we again get the results just listed when we study the effect of more IG effort on the newsvendor's average orders. But there are certainly IGs for whom the above results are sharply violated.

Consider the following simple example. Demand is $0, 1$ or 2 with respective prior probabilities $.1, .1, .8$, so expected demand is 1.7. The IG has three effort levels.

- At the lowest effort, the IG provides no new information and the newsvendor's posterior probabilities are the same as the priors.

- At the intermediate effort level, the IG sends the newsvendor one of two signals denoted y_1 and y_2. Signal y_1 tells the newsvendor that demand is 0 with probability $\frac{1}{2}$ and 1 with probability $\frac{1}{2}$. Signal y_2 tells the newsvendor that demand is certainly 2. The signal probabilities are $.2$ for y_1 and $.8$ for y_2.

- At the highest effort level, the IG provides the newsvendor with perfect information about demand.

For a given value of c in $(0, 1)$, it is straightforward to show that expected newsvendor profit rises when we go from lowest IG effort to intermediate effort and from intermediate effort to highest effort, and to calculate, for each effort level, the newsvendor's average optimal *ex ante* order quantity and average *ex post* sales. (For intermediate effort, we average over the two possible signals). Figure 1 shows the relation between the cost c and the average order quantity for each of the three effort levels.

Notice that the figure differs from our result [**2′**]. For a fixed value of c in $(.5, .8)$, the average order quantity drops when the IG goes from lowest effort to intermediate effort, but rises when the IG goes from intermediate effort to highest effort. That means that [**2**] is also violated: there is no critical value of cost such that for all costs on one side, the average order rises whenever the IG works harder and for all costs on the other side, the average order falls whenever the IG works harder. When we turn to our scale/location IG, we will obtain two figures (Figures 4 and 5 below) which again show the relation between c and average order quantity. They will obey statements [**2**] and [**2′**] and will contrast sharply with Figure 1. (Figure 4 will portray a uniform-distribution example and Figure 5 will portray a normal-distribution example).

As for average *ex post* sales, we find, in our three-demand example, that statement [**3**] is also violated: for $.5 < c < .8$, average *ex post* sales are 1.7 for lowest effort, 1.6 for intermediate effort, and 1.7 again for highest effort.

A remark on the Blackwell IG. One might conjecture that results about the effect of more IG effort on the newsvendor's average orders are easily obtained if we let our IG fit the assumptions of the celebrated papers by Blackwell on comparisons of experiments. (See Blackwell (1951), (1953); detailed treatments are found in Marschak and Miyasawa (1968), in Chapter 1 of Marschak and Radner (1972), and in Marschak, Shanthikumar, and Zhou (2013)). The Blackwell theorems have played a large role in the economics of information. Consider an IG whose signals are received by an action-taker. The action-taker responds to a signal by choosing an action which maximizes expected payoff for the posterior implied by the signal. Now consider two of the IG's effort levels and suppose that *whatever the action-taker's payoff function may be, and whatever the prior may be*, the action-taker benefits (weakly) when the IG switches from the lower effort to the higher effort: after the switch, the best expected payoff, averaged over the IG's signals, rises or stays the same. Let us call such an IG a *Blackwell IG* with respect to the two efforts. One example of a Blackwell IG is an IG who partitions the states of the world and tells the action-taker the set in which the true state lies. The IG's second effort refines the first effort's partitioning. That refinement cannot harm the action-taker, no matter what the payoff function and the prior may be. In a second example, the states are the proportion of persons that are of the first type in a two-type population, and the action-taker's payoff depends on the action he chooses and on the true proportion of type-1 persons. (We suggested such an example above for the case where the action-taker is a newsvendor with two types of demanders). The Blackwell IG samples the population, learns each sampled person's true type, and reports the result to the action-taker. A larger sample — i.e., more IG effort — can never hurt the action taker, no matter what the payoff function and the prior may be.

The comparison-of-experiments literature shows (among many other results) that an IG is a Blackwell IG with respect to the two efforts if and only if the following holds: for any convex function ϕ on the possible posteriors, the expected value of ϕ, over the IG's signals, is at least as high for the larger effort as for the smaller effort. If the action-taker is our newsvendor, and if his best order quantity were a convex (concave) function of the posterior, then we could immediately
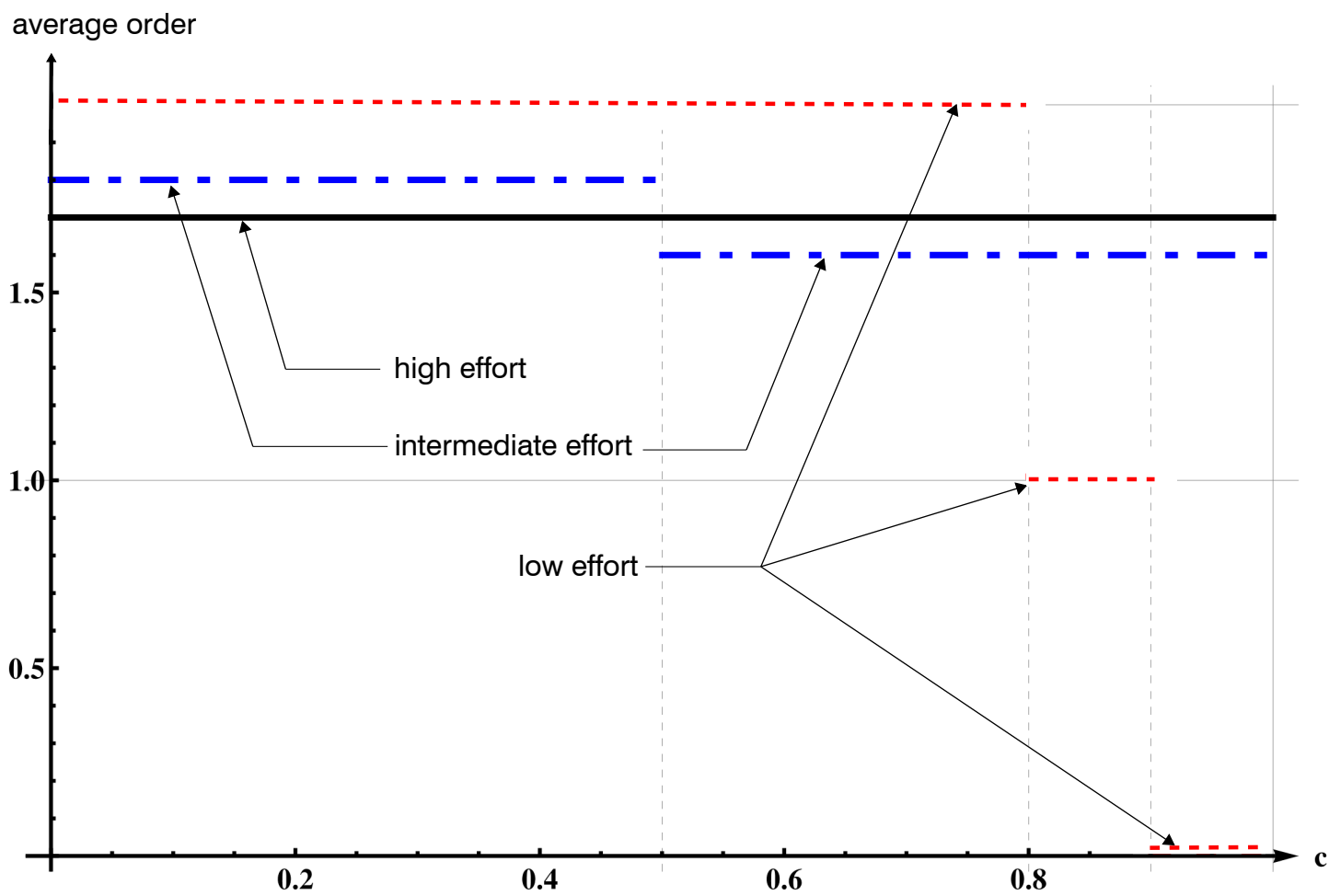
4

Figure 1

average order: the three-demand example

conclude that his average order cannot drop (cannot rise) when the Blackwell IG makes an effort switch that benefits the newsvendor. Unfortunately the required convexity (concavity) does not, in general, hold. Moreover, our scale/location IG, for whom effort is defined by the average of the scale parameters, is *not*, in general, a Blackwell IG with respect to the two efforts. For some scale/location IGs we can find a prior and a payoff function such that lower average scale implies a lower average value of the action-taker's best expected payoff, and another prior and payoff function such that lower average scale implies a higher average value of the action-taker's best expected payoff. In summary, the Blackwell theorems do *not* provide a shortcut in our investigation.

Some related literature.

We shall discuss four groups of papers.

First, many papers have studied the effect of changes in the distribution of demand on an inventory manager's profit and order size but are not directly concerned with the improvement of demand information. An early paper in this group is Gerchak and Mossman (1992), where demand (at a fixed price) is a weighted sum of a fixed component and a random component. Higher uncertainty means that more weight is given to the latter. The effect of more uncertainty on the profit and order size for a single-period newsvendor is studied. Ridder *et al* (1998) study a single-period newsvendor and compare the inventory cost under two demand distributions that are ordered according to various forms of n-th order stochastic dominance. They show that under second order stochastic dominance, lower demand variance always implies lower cost. But they give conditions for third or higher order stochastic dominance under which inventory cost can be lower while the variance is higher. Papers by Song and Zipkin (1993), Song (1994), and Song *et al* (Song, Zhou, Hou, Wang) (2010) have an analogous agenda in a dynamic inventory setting. They trace the effect of variability in lead time (time until the ordered inventory is delivered) on inventory cost. In Aggrawal and Seshadri (2000), Li and Atkins (2005), and Xu, Chen, and Xu (2010) both price and quantity are chosen, the effect of price on demand is random, and the variability of that effect influences the inventory manager's profit and order size.

In a second group of papers the inventory manager's information about future demand can be improved by advance ordering. If a buyer commits to an order to be delivered at a future date, that removes the uncertainty about that buyer's future demand. Moreover the advance order may provide clues as to the future demand of those buyers who do not order in advance. An early paper in this group is Buzacott and Shanthikumar (1994), which investigates the value of better forecasts obtained by advance demand information in production/inventory systems. It is shown that the improved forecast serves as a substitute for safety stock. Another early paper is Hariharan and Zipkin (1995). They study a dynamic inventory system where customers randomly arrive and each places an order for delivery at a later time. A delivery later than the time the customer selected may occur. The main conclusion is that when an optimal base-stock policy is followed, an increase in average demand lead time (the length of time the customer agrees to wait) has the same effect on system performance as a drop in average supply lead time (the time between receipt of the order by the supplier and its delivery).

Another early contribution to the advance demand information literature is Gallego and Özer (2001). They find optimal state-dependent policies for stochastic inventory systems, where the state reflects the current advance demand information. The benefits of advance demand information are investigated in further detail in Özer (2003).In Lu, Song, and Yao (2002) a similar agenda is pursued (in the paper's section 6) for the more complex setting of a multicomponent assemble-to-order system. It is now found that knowing demand in advance is more useful (there are fewer late deliveries) than reducing the supply lead time for the components. Papers by Özer and Wei (2004) and by Karaesmen *et al* (2004) study capacitated production systems and find that advance demand information can be a substitute for both capacity and inventory. Other papers reaching similar conclusions include Gayon, Benjaafar and de Vericourt (2009) and Karaesmen, Buzacott, and Dallery (2002). Song and Zipkin (2012) also obtain similar results for a model in which demand is driven by underlying random variables like weather or the behavior of competitors, and advance information about those variables is revealed by advance orders. Later papers (Boyaci and Özer (2010), Prasad, Stecke, and Zhao (2011)) extend the problem to include the inducement of advance demand by appropriate pricing policies. If discounts are used, the cost of the discount has to be balanced against the value of the improved demand information.

A third group of papers consider the sharing of demand information. In Lee, So, and Tang (2000) a retailer's demands follow a process known to the retailer. In each period, the retailer has an inventory, observes current demand, and orders from the manufacturer an amount that permits the retailer to meet current demand and replenish inventory. The manufacturer, in turn, anticipates the retailer's orders, maintains his own inventory, and places his own replenishment orders with a supplier. The paper studies the benefits of information sharing by comparing the case where the manufacturer does not know the retailer's demand process and the case where he does (sharing). The comparison is made with respect to the manufacturer's expected cost and expected inventory. It is found, under general conditions, that both of these are lower under sharing. A similar conclusion is reached in Cachon and Fisher (2000), where there are several retailers and they face a stationary demand process. The manufacturer either knows the retailers' orders or knows both their orders and their inventory positions as well (sharing). In Gaur, Giloni, and Seshadri (2005) the problem posed in Lee, Song, and Tang is revisited for the case where retailer demand follows an autoregressive moving average process. It is found that whenever the manufacturer benefits from sharing, the manufacturer's average safety stock declines. Many papers focus on strategic aspects of information sharing. They study a game played by various members of a supply chain. The players may be a manufacturer and retailers, as in Zhang (2002) and Li and Zhang (2008). In an equilibrium of the game, the retailers may choose to convey private information about demand and inventory positions to the manufacturer, who uses it in choosing a price. These papers do not directly address the effect of sharing on inventory size, i.e., the average inventory size when sharing occurs at the game's equilibrium and the size when it does not.

Finally, there are papers which do not fall easily into the three preceding groups but nevertheless investigate, in various other ways, the effect of improved information about demand on profit and on inventory size. An early paper is Milgrom and Roberts (1988), in which a producer

chooses a collection of varieties of its product and each customer demands one variety. The producer can stock a quantity of each variety, or he can survey (at a cost) a fraction of customers, obtaining perfect information for each surveyed customer. It is found that the producer does one or the other, but not both. Another early paper is Dudley and Lasserre (1989). Here demand is determined by a stochastic process and a forecast of next period's demand is based on observed demand in a past period. If the past period is more recent, the forecast costs more but it is more reliable. A producer chooses his forecast expenditure and responds to the forecast by choosing inventory size. It is found that when there is a drop in the cost of a forecast (for every given past period), then the average inventory size drops. Cachon and Fisher (1997), a paper based on observed practice, studies an inventory management system in which each retailer's daily demands are used to make an exponential-smoothing forecast of future demand. A simple but practical reorder rule uses the forecasts. Simulations suggest that the forecasts reduce excess inventories. Bergen and Iyer (1997) consider the improvement in demand forecasts when we shorten the time between (1) the retailer's placement of an order and (2) production and delivery of the order by the manufacturer. Final demand becomes known very shortly after (2) occurs. Hence the time between (1) and the realization of demand is reduced. Forthcoming demand is better estimated when it is just a short time away. So there is a tradeoff between the extra cost of rapid production and delivery on the one hand and improved inventory decisions due to better forecasts on the other. The paper studies the preferred tradeoff from the viewpoint of the manufacturer and that of the retailer. Among other results, one finds that shortening the lead time reduces the retailer's average order. That occurs as well in a later paper by A.H. Lau and H.S. Lau (2000), where some of the assumptions in Bergen and Iyer are relaxed. Gurnami and Tang (1999) study a retailer who can place an order with a supplier at one of two time points. Both points occur before demand is realized. Information about the forthcoming demand is summarized by a one-dimensional signal. As time passes, more becomes known about the forthcoming demand, so the correlation between the signal and the true demand is higher at the later point than at the earlier one. However, unit cost is random (as well as demand) and its average is higher for orders placed later. Under certain conditions on the random cost, it is found that when the retailer behaves optimally, his average early order drops when the correlation increases. There is a similar result in Choi, Li, and Yu (2003), where a related two-time-point model with Bayesian updating is studied.

Toktay and Wein (2001) study a production manager who obtains demand forecasts and chooses a production plan that meets demand while minimizing expected inventory costs. Demand follows a stationary stochastic process. The forecaster may specify the process incorrectly. The effect of improved forecast quality is studied. Under certain conditions, higher quality leads to smaller inventories. In Jain and Moinzadeh (2005) there is uncertainty about the manufacturer's available supply and better information about the supply allows the retailer to have a smaller safety stock.

Iyer, Narasimhan, and Niraj (2007) consider a retailer, a manufacturer, and an information system which provides signals about demand. Demand has two possible values and there are two signals. The information system's reliability is measured by a number in $[0, 1]$, namely the probability of high demand given the signal called "high", which is assumed to equal the

probability of low demand given the signal called "low". "More information" means a higher value of that number. If first-best decisions are made, more information leads to lower excess stock.

Bensoussan, *et al* (2009) is a recent paper which explicitly addresses the question whether demand information is a substitute for inventory size, and finds that this is indeed the case. There is a single decision-maker, an "Inventory Manager", who confronts a (finite) sequence of demands and chooses an order in each period as a function of what he knows about the demand history thus far. The most recently known demand may be some periods earlier. The delay can be reduced. The manager's cost in each period demands on the period's starting inventory and on the new order. Unmet demand is backlogged and fulfilled in the next period. Each period's order maximizes discounted future cost. For a given delay, the dynamic-programming solution is characterized. Shorter delays indeed lead to smaller average inventories.

On the other hand, Anand (2009) presents an innovative infinite-period model in which the conclusion is quite different. A producer and an inventory manager seek to maximize discounted expected profit. In each period, the inventory manager observes a signal about the forthcoming demand. A myopic policy, which is shown to be optimal, determines the current period's production quantity, the manager's current order, and the amount the manager ships to demanders. The manager's shipment is the smaller of his current inventory and a "ship-up-to" threshold. The threshold depends on the current signal. The signal is supplied by what we have called, in the preceding discussion, a Blackwell IG. It is found that if the ship-up-to threshold is concave in the signal (which fits a variety of examples), then we have "complements" rather than "substitutes": an increase in the Blackwell IG's effort leads to higher average inventories.

The path we follow here appears to be new: choose a model of an IG, define effort as a function of the IG's signal *set*, and investigate the effect of higher effort on the newsvendor's average orders and average *ex post* sales in a single-period setting.

Plan of the rest of the paper. In Section 2 we formally describe our Information-gatherer, who performs a scale/location transform of a base distribution. We discuss two examples of such an IG in detail: a normal-distribution example and a uniform-distribution example. We briefly sketch two other examples: a triangular-distribution example and a finite example. Section 3 identifies a key property of the newsvendor payoff function: it is shift-invariant and homogeneous of degree one. Section 3 then presents two key lemmata which exploit that key property as well as the scale/location property of our IG. Section 4 presents two propositions about the usefulness of more IG effort for the classic newsvendor and the effect of more effort on the average *ex ante* order and the average *ex post* sales. In Section 5 we consider a generalized "regret-averse" newsvendor and we obtain two propositions that are the analogs of the propositions in Section 4. Section 6 provides some concluding remarks and points to a large unexplored terrain for future research.

2. The IG who transforms a base distribution and sends a signal to the newsvendor.

The true demands lie in a set $\mathcal{D} \subseteq I\!\!R$. Once true demand has been chosen by "Nature", the IG observes a signal y which is conveyed to the newsvendor. Our IG/newsvendor model has the following elements:

- The newsvendor and the IG know the joint distribution of the pair (D, y), The marginal distribution of D is the prior on $\mathcal{D}$.

- For a given y, the joint distribution determines the posterior CDF on the true demands; once he receives the signal y, the newsvendor uses that posterior to calculate his order.

- For every y, the posterior CDF has the following property in our model: it is a scale/location transform of a fixed *base* CDF, denoted B, which has mean zero. The location parameter is denoted $a(y)$. The scale parameter is positive and is denoted $b(y)$. The posterior CDF (a transform of B) is denoted $T_{a(y),b(y)}$.

Let z denote the base random variable. Since its mean is zero, the mean of the transformed variable $D = a(y) + b(y)z$ is $a(y)$. The variance of D is $(b(y))^2$ times the variance of the base random variable. Note that the base CDF B and the transformed CDF $T_{a(y),b(y)}$ — which is the newsvendor's posterior — obey the following:

$$T_{a(y),b(y)} \ (D) = B\left(\frac{D - a(y)}{b(y)}\right).$$

That fact will be used in several of our proofs. The IG has a collection of possible signals. Each signal identifies a location/scale pair and hence a posterior CDF. We shall call the collection of posterior CDFs a *structure*. We will want to compare alternative structures. To do so, we use an index η. A value of η identifies a particular structure. For that structure, the set of possible signals is denoted Y^η. For each signal y in Y^η, the location/scale pair is denoted $(a(y,\eta), b(y,\eta))$ and the corresponding transform of the base B is $T_{a(y,\eta),b(y,\eta)}$. That transform is the newsvendor's posterior CDF. It is helpful to use an abbreviated symbol for the newsvendor's CDF, namely F_y^η. Thus

$$F_y^\eta = T_{a(y,\eta),b(y,\eta)}.$$

For a given η, we shall call the pair

$$I_\eta = \langle Y^\eta, \{F_y^\eta\}_{y \in Y^\eta}\rangle$$

the *information structure identified by* η.

The symbol E_{F_y} will denote expectation under the CDF F_y, and the symbol $E_{y \in Y}$ will denotes expectation with respect to signals in the set Y. We shall say that the family of structures $\{I_\eta\}$ has the *average-scale-reduction order* or, for brevity, the *scale-reduction order*, if

$$E_{y \in Y^{\eta'}} \ b(y, \eta') < E_{y \in Y^\eta} \ b(y, \eta) \text{ whenever } \eta' > \eta.$$

If a family of structures has the scale reduction order, then it is reasonable to take η as an IG-effort index: a higher value of η means a drop in the average, over all signals, of the transformed distribution's standard deviation. We shall show that scale reduction is useful to the newsvendor: it increases the average, over all signals, of his highest attainable expected profit given the signal. So for a family that has the scale reduction order, the first of two structures is more useful than

the second if and only if the first requires more IG effort. That is our statement [**1**] in the Introduction. Moreover, we shall find, as in statements [**2**] and [**3**], that scale reduction (more effort) always leads to higher average *ex post* sales, but for certain values of the cost c it leads to a higher average *ex ante* order, and for other values to a lower average *ex ante* order.

Note that the example in the Introduction, where demand D takes just three values, does *not* have the location/scale property. Consider the two signals in the intermediate-effort structure. For the first signal, the posterior assigns positive probability to the first two demands only. For the second signal, the posterior assigns positive probability to the third demand only. There does not exist a base distribution B on the three demands such that each of the two posteriors is a scale/location transform of B.

2.1 First example: the IG has a uniform base distribution and partitions a uniformly distributed set of true demands

Suppose that the set of possible true demands is $\mathcal{D} = [0,1]$ and the prior is the uniform distribution on $[0,1]$. Let the IG's base distribution B be the uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$. For the information structure I_η, there are n_η possible signals and the signal set is $Y^\eta = \{y_1(\eta), y_2(\eta), \cdots, y_{n_\eta}(\eta)\}$. The typical signal $y \in Y^\eta$ defines the location parameter $a(y, \eta)$ and the scale parameter $b(y, \eta)$. These, in turn, define a transform of B. The transform is the newsvendor's posterior CDF given the signal y. The posterior is uniform on the interval $[a(y,\eta) - \frac{b(y,\eta)}{2}, a(y,\eta) + \frac{b(y,\eta)}{2}] \subset [0,1]$. The width of the interval equals $b(y,\eta)$. The probability of the signal y equals that width divided by the width of $\mathcal{D} = [0,1]$, i.e., it equals $b(y,\eta)$. So the average scale parameter for the structure I_η is

$$E_{y \in Y^\eta}\, b(y, \eta) = \sum_{i=1}^{n_\eta} \left[b\Big(y_i(\eta), \eta\Big) \right]^2.$$

The IG's structure family $\{I_\eta\}$ has the scale-reduction order (i.e., $E_{y \in Y^{\eta'}}\, b(y) < E_{y \in Y^\eta}\, b(y)$ whenever $\eta' > \eta$) if

$$\sum_{i=1}^{n_\eta} \left[b\Big(y_i(\eta), \eta\Big) \right]^2 \text{ is strictly decreasing in } \eta.$$

That is the case, in particular, if the IG always partitions $[0,1]$ into *equal* intervals and there are more of them when η rises, i.e., $b(y_i(\eta)), \eta) = \frac{1}{n_\eta}$ for all i, where n_η is strictly increasing in η. Higher η then means that, on the average, the posterior distribution has narrower support.

To illustrate further, let the IG's family of possible structures contain just two structures, I_η and $I_{\eta'}$. The structure I_η partitions $[0,1]$ into just two intervals, $[0, \frac{1}{5}]$ and $(\frac{1}{5}, 1]$. The structure has two signals. Signal, $y_1(\eta)$ —or y_1, for brevity — identifies the location/scale pair $(a(y_1, \eta), b(y_1, \eta)) = \left(\frac{1}{10}, \frac{1}{5} \right)$ and signal y_2 identifies the pair $(a(y_2, \eta), b(y_2, \eta)) = \left(\frac{3}{5}, \frac{4}{5} \right)$. The first pair tells the newsvendor that demand is uniformly distributed on $\left[a(y_1, \eta) - \frac{b(y_1,\eta)}{2}, a(y_1, \eta) + \frac{b(y_1,\eta)}{2} \right] = [0, \frac{1}{5}]$. The second pair tells him that demand is uniformly distributed on $\left[a(y_2, \eta) - \frac{b(y_2,\eta)}{2}, \right.$

10

$a(y_2, \eta) + \frac{b(y_2,\eta)}{2}\bigg] = \left[\frac{1}{5}, 1\right]$. The two signal probabilities are the two subinterval widths $\frac{1}{5}, \frac{4}{5}$, and the average scale parameter is $\left(\frac{1}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = \frac{17}{25}$.

Structure $I_{\eta'}$, on the other hand, partitions $[0, 1]$ into the three intervals $[0, \frac{1}{4}], (\frac{1}{4}, \frac{3}{4}], (\frac{3}{4}, 1]$. It has three signals. They identify, respectively, the location/scale pairs $\left(\frac{1}{8}, \frac{1}{4}\right), \left(\frac{1}{2}, \frac{1}{2}\right)$, and $\left(\frac{7}{8}, \frac{1}{4}\right)$, and those pairs tell the newsvendor, respectively, that demand is uniformly distributed on $[0, \frac{1}{4}]$, on $[\frac{1}{4}, \frac{3}{4}]$, and on $[\frac{3}{4}, 1]$. The three signal probabilities are the three subinterval widths $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$, and the average scale parameter is $\left(\frac{1}{4}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{4}\right)^2 = \frac{3}{8}$. That is smaller than $\frac{17}{25}$, so the IG's effort increases when he goes from structure I_η to structure $I_{\eta'}$. We shall return to the two structures $I_\eta, I_{\eta'}$ below, when we illustrate our general Propositions.

2.2 Second example: demand is normally distributed and the IG observes demand plus a normally distributed noise.

In the next example the set $\mathcal{D}$ of demands is the real line. The prior is normal with mean μ and variance $\frac{1}{\zeta}$, where $\zeta > 0$. The IG observes demand plus a normally distributed noise. For a given true demand D we have the following for the signal y:

$$y = D + \epsilon; \ \epsilon \text{ is normally distributed with mean zero and variance } \frac{1}{\eta}.$$

The random pair (D, y) has a bivariate normal distribution. Each of its two means equals μ. The two variances are $\frac{1}{\zeta}$ and $\frac{1}{\zeta} + \frac{1}{\eta}$. The covariance equals $\frac{1}{\zeta}$. We find that the conditional distribution of D given a signal y is normal with mean

$$a(y, \eta) = \frac{\eta y + \zeta \mu}{\eta + \zeta}$$

and standard deviation

$$b(y, \eta) = \sqrt{\frac{1}{\eta + \zeta}}.$$

We can now describe the IG's procedure as a scale/location transformation, where the above η is the IG's index. The base distribution is the standard normal $\mathcal{N}(0, 1)$. The IG's set of possible indices η is the positive real line. For every η, the signal set Y^η is the real line. When the newsvendor receives the signal $y \in Y^\eta$, he replaces the prior with the posterior F_y^η, namely the normal CDF with mean $a(y, \eta)$ and variance $(b(y, \eta))^2$.

Since the term y does not enter the expression for $b(y, \eta)$, it is automatically the case that if $\eta' > \eta$ then we have

$$E_{y \in Y^{\eta'}} \ b(y, \eta') = \sqrt{\frac{1}{\eta' + \zeta}} < \sqrt{\frac{1}{\eta + \zeta}} = E_{y \in Y^\eta} \ b(y, \eta).$$

So the family of structures I_η has the scale reduction order. When the IG switches from η to $\eta' > \eta$, he works harder. The average standard deviation of noise is reduced. For a given confidence level, the average width of a confidence interval around the mean for the posterior

11

obtained in $I_{\eta'}$ (averaging over the signals in $Y^{\eta'}$) is less than the average width for the posterior obtained in I_η (averaging over the signals in Y^η).

2.3 Other examples: triangular and finite.

Consider the case where the prior distribution of demand is symmetric and triangular with support $[0, 1]$. Let the IG's base density function be symmetric and triangular with support $\left[-\frac{1}{2}, \frac{1}{2}\right]$. In a location/scale transform of the base density, where the location/scale parameter pair is (a, b), we shift the apex of the base triangle to the right by the distance a (i.e., we add a to the mean), and we multiply the width of the base support by b. We then have another symmetric triangular density function, where the coordinates of the apex are $\left(a, \frac{2}{b}\right)$, the mean is a, and the support is $\left[a - \frac{b}{2}, a + \frac{b}{2}\right] \subseteq [0, 1]$. That transformed density function is the newsvendor's posterior density function when the signal he receives identifies the pair (a, b). Now consider an IG with two structures. In each structure, a signal identifies a location/scale pair. Effort is higher in the second structure if the average scale parameter, over all the signals, is lower. In the first structure, the IG has *three* signals, with probabilities $\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$. For the first signal, $(a, b) = \left(\frac{1}{4}, \frac{1}{2}\right)$, so the posterior density function is symmetric and triangular with mean $\frac{1}{4}$ and support $\left[0, \frac{1}{2}\right]$. For the second and third signals, the (a, b) pairs are $\left(\frac{1}{2}, \frac{1}{2}\right)$ and $\left(\frac{3}{4}, \frac{1}{2}\right)$. Those define posterior symmetric triangles with supports $\left[\frac{1}{4}, \frac{3}{4}\right]$ and $\left[\frac{1}{2}, 1\right]$. The average scale parameter is $\frac{1}{2}$. In the second structure, the IG has *five* signals, with probabilities $\left(\frac{1}{16}, \frac{1}{8}, \frac{1}{16}, \frac{1}{2}, \frac{1}{4}\right)$. The (a, b) pairs are $\left(\frac{1}{8}, \frac{1}{4}\right), \left(\frac{1}{4}, \frac{1}{4}\right), \left(\frac{3}{8}, \frac{1}{4}\right), \left(\frac{1}{2}, \frac{1}{2}\right)$, and $\left(\frac{3}{4}, \frac{1}{2}\right)$. The average scale parameter is now $\frac{1}{4} \cdot \frac{1}{16} + \frac{1}{4} \cdot \frac{1}{8} + \frac{1}{4} \cdot \frac{1}{16} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = \frac{7}{16}$, so the second structure indeed requires more effort than the first.

For a simple finite example, let the possible demands be $D^*, D^* + L, D^* + 2L, D^* + 3L, D^* + 4L, D^* + 5L$, where $L > 0$. In the prior distribution each of these has probability $\frac{1}{6}$. Let the base variable take just three values, namely $-1, 0, 1$, each with probability $\frac{1}{3}$. Note that the mean of the base distribution is zero and that the three mass points of the support $\{-1, 0, 1\}$ are evenly spaced with a gap of *one* between the successive points. In an (a, b) location/scale transform of the base distribution we again have three evenly spaced mass points in the support. The new mean is a plus the base distribution's mean, i.e. it is $a + 0 = a$. The gap between the successive mass points is now b times the base distribution's gap, i.e., it is $b \cdot 1 = b$. Now consider an IG with three structures. The first structure has *two* signals, each with probability $\frac{1}{2}$. The first signal tells the newsvendor that the support of the demand distribution is $\{D^*, D^* + 2L, D^* + 4L\}$, where the three probabilities are equal. The posterior is then an (a, b) transform of the base, where $(a, b) = (D^* + 2L, 2L)$. The second signal tells the newsvendor that the support is $\{D^* + L, D^* + 3L, D^* + 5L\}$ (with equal probabilities). This time $(a, b) = (D^* + 3L, 2L)$. The second structure also has *two* signals, each with probability $\frac{1}{2}$. The first signal tells the newsvendor that the support is $\{D^*, D^* + L, D^* + 2L\}$ (with equal probabilities), so $(a, b) = (D^* + L, L)$. The second signal tells the newsvendor that the support is $\{D^* + 3L, D^* + 4L, D^* + 5L\}$ (with equal probabilities), so $(a, b) = (D^* + 4L, L)$. Finally, the third structure has *four* signals, each with probability $\frac{1}{4}$. The first two signals provide the same information as the two signals of the first structure and they have the same (a, b) pairs as the first structure's signals. The last two signals provide the same information as the two signals of the second structure and they have the same

(a, b) pairs as the second structure's signals. Average scale is $2L$ for the first structure, L for the second structure, and $\frac{3}{2}L$ for the third structure. So effort rises when the IG goes from the first structure to the third and then from the third to the second.

We now prove general propositions about the effects of more IG effort (lower average scale) on the newsvendor. The propositions apply, in particular, to the four examples just discussed.

3. A property of the newsvendor's payoff function and two key lemmata.

The newsvendor's payoff function

$$u(q, D) = \min(D, q) - cq$$

has the following property:

(1)
$$\begin{cases} \text{There exists a number } \gamma \text{ such that for any } \alpha \text{ and any } \beta > 0, \\[2mm] u(q, D) = \gamma\alpha + \beta \cdot u\left(\frac{q - \alpha}{\beta}, \frac{D - \alpha}{\beta}\right) \text{ for all } q, D. \end{cases}$$

So the function u is shift-invariant and homogeneous of degree one. The role of γ is played by $1 - c$.

To obtain results about the effect of a change in our IG's structure on the newsvendor's order quantities we need two key lemmata. Lemma A exploits the shift invariance and homogeneity of the newsvendor's payoff function. Those properties articulate well with the scale/location property of our IG's family of structures. That allows us to obtain a key linear relation between the average best quantity if demand had the base distribution and the average best quantity when demand has the transformed distribution. Lemma B obtains a similar relation between the base distribution's highest attainable expected profit and the highest attainable expected profit for the transformed distribution.

The proof of each lemma, and the proofs of each of the subsequent propositions, are found in the Appendix.

Lemma A

Let $u = \min(q, D) - cq$ be the newsvendor's payoff function. Consider an IG who has a base CDF B with mean zero. The IG's signal determines a location parameter a and a scale parameter $b > 0$. They define a transformed CDF T_{ab}. Let q_B denote the smallest maximizer of $E_B \, u(q, D)$ and let q_{ab} denote the smallest maximizer of $E_{T_{ab}} \, u(q, D)$. Let u_B denote $E_B \, u(q_B, D)$ and let $u_{T_{ab}}$ denote $E_{T_{ab}} \, u(q_{ab}, D)$. Then

(A1)
$$q_{ab} = a + b \cdot q_B.$$

and

(A2)
$$u_{T_{ab}} = \gamma a + bu_B, \text{ where } \gamma = 1 - c.$$

13

When there are several maximizers of $E_B\, u(q, D)$, the proof requires us to identify one of them. Our proof uses the smallest maximizer.

To state the second lemma, we let $U(\eta)$ denote the average, over all the signals in Y^η, of the newsvendor's expected value of the profit u when he responds to each signal by maximizing expected profit under the posterior (the transformed distribution) that the signal identifies. The symbols q_B, u_B, as well as symbols of the form q_{ab}, all have the same meaning as in Lemma A. Each signal y in Y^η identifies a location parameter $a(y, \eta)$ and a scale parameter $b(y, \eta)$. We let $\hat{q}(\eta)$ denote the average, over all the signals in Y^η, of the newsvendor's smallest maximizing quantity. Thus

$$\hat{q}(\eta) \equiv E_{y \in Y^\eta}\, q_{a(y,\eta),b(y,\eta)}; \;\; U(\eta) \equiv E_{y \in Y^\eta}\, \left(u\left(q_{a(y,\eta),b(y,\eta)}, D\right)\right).$$

We let $V(\eta)$ denote the average of the scale parameter, over all the signals in Y^η. So

$$V(\eta) \equiv E_{y \in Y^\eta}\, b(y, \eta).$$

A structure family $\{I_\eta\}$ has the scale reduction order if $V(\eta') < V(\eta)$ whenever $\eta' > \eta$.

Lemma B

For every pair (η', η), we have

(B1)
$$\hat{q}(\eta') - \hat{q}(\eta) = q_B \cdot (V(\eta') - V(\eta))$$

(B2)
$$U(\eta') - U(\eta) = u_B \cdot (V(\eta') - V(\eta)).$$

Note that (B1) and (B2) concern, respectively, the change in expected order quantity and the change in expected profit when we move from one structure to another. We find that the first change equals a constant times the change in average scale, and the second change equals another constant times the change in average scale. The first constant is the base distribution's average order and the second constant is the base distribution's average profit.

4. Two propositions about the change in the classic newsvendor's order quantity when the IG switches from one structure to another.

In the first of the two propositions which now follow, we consider the classic newsvendor's *ex ante* order quantity. Recall that an IG has a collection of possible signals, each identifying a posterior. We have called the collection a *structure*. We now consider a switch from one structure to another. We find that a switch in the IG's structure is useful to the newsvendor —

14

it increases the average profit U — if and only if the switch also reduces V, the average scale parameter. We also find a critical value of the cost c. On one side of the critical value, a useful switch increases the average order. On the other side, such a switch decreases the average order.

Proposition 1

Consider a classic newsvendor whose profit is $\min(q, D) - cq$, with $0 < c < 1$. Consider the scale/location IG, who has a base CDF B with mean zero and a density that is positive at every point of its support. Let $I_{\eta'}$ and I_{η} be two of the IG's information structures. Then the following statements hold.

[1] $V(\eta') < V(\eta)$ if and only if $U(\eta') > U(\eta)$.

[2] If $1 - c < B(0)$, then $U(\eta') > U(\eta)$ if and only if $\hat{q}(\eta') > \hat{q}(\eta)$.

[3] If $1 - c > B(0)$, then $U(\eta') > U(\eta)$ if and only if $\hat{q}(\eta') < \hat{q}(\eta)$.

[4] If $1 - c = B(0)$, then $\hat{q}(\eta') = \hat{q}(\eta)$.

To interpret the number $B(0)$, recall that the base CDF B has mean zero. If the base CDF has a symmetric density function, then $B(0) = \frac{1}{2}$. That is the case in the uniform example in 2.1 and the normal example in 2.2. So in those examples the critical value of c is $\frac{1}{2}$. Note that for the normal-distribution case there is a direct way of seeing the criticality of $c = \frac{1}{2}$. The direct argument uses the fact that when demand is normal, the optimal order equals mean demand plus standard deviation times the order that is optimal when mean is zero and standard deviation is one. Proposition 1 uses a more general argument to obtain the criticality of $c = \frac{1}{2}$, and that argument covers the normal case, the uniform case, and the triangular case discussed in 2.3. It turns out that in all these cases the criticality of $c = \frac{1}{2}$ is explained by the fact that every posterior is a scale/location transform of a common base distribution. In Proposition 3 we will obtain a generalization of the criticality of $c = \frac{1}{2}$ for a generalized newsvendor. There, however, we see no direct path leading to that generalized conclusion in the normal-distribution case.

Illustrating Proposition 1 in a uniform-distribution example. Consider the uniform-distribution example in **2.1** above. Demand is distributed uniformly on $[0, 1]$ and that is the newsvendor's prior. Consider the two structures, I_{η} and $I_{\eta'}$. In I_{η}, the IG partitions $[0, 1]$ into n subintervals having probabilities (widths) $p_1, \ldots, p_n$, where $\sum_{1=1}^{n} p_i = 1$. The newsvendor learns the subinterval that contains the current demand. It is straightforward to show that

$$(2) \qquad \text{average profit over all subintervals } = \frac{c \cdot (c - 1)}{2} \cdot \sum_{i=1}^{n} (p_i)^2 + \frac{1}{2}.$$

In $I_{\eta'}$ the IG partitions $[0, 1]$ into k subintervals having probabilities (widths) $r_1, \ldots, r_k$, where $\sum_{1=1}^{k} r_i = 1$. For I_{η}, the average scale parameter is $\sum_{i=1}^{n} (p_i)^2$ and for $I_{\eta'}$, it is $\sum_{i=1}^{k} (r_i)^2$.

15

The structure $I_{\eta'}$ requires more scale-reduction effort than I_η if

$$\sum_{1=1}^{k} (r_i)^2 < \sum_{1=1}^{n} (p_i)^2 .$$

Suppose, in particular, that I_η and $I_{\eta'}$ are the two specific structures presented in section 2.1 above. In I_η there are two subintervals, $[0, \frac{1}{5}]$ and $(\frac{1}{5}, 1]$. In the higher-effort structure $I_{\eta'}$ there are three subintervals $[0, \frac{1}{4}], (\frac{1}{4}, \frac{3}{4}], (\frac{3}{4}, 1]$. In Figure 2 we illustrate statement [1] of the Proposition by using statement (2) and graphing the average profits $U(\eta)$ and $U(\eta')$ for all values of the unit cost c. The figure has two additional average-profit graphs. One concerns perfect information, where the newsvendor knows demand exactly and his order always matches demand. The other concerns the no-information case, where the newsvendor only knows that demand is uniform on $[0, 1]$.

$\boxed{\text{FIGURE 2 HERE}}$

Next we illustrate statements [2]-[4] of the Proposition, using the same specific structures I_η and $I_{\eta'}$ as in Figure 2. Figure 3 shows the relation between the unit cost c and the average orders (over all signals), $\hat{q}(\eta)$ and $\hat{q}(\eta')$. When I_η has n subintervals with probabilities (widths) $p_1, \ldots, p_n$, it is straightforward to establish that

(3)
$$\hat{q}(\eta) = \left(\frac{1}{2} - c\right) \cdot \sum_{i=1}^{n} (p_i)^2 + \frac{1}{2}.$$

Using (3), we calculate average order for the structures $I_\eta, I_{\eta'}$ of Figure 2. For each structure, we graph average order as a function of c in Figure 3. The figure again has two further average-profit graphs. One concerns perfect information, where average order is $\frac{1}{2}$, the mean of the true demands. The other concerns the no-information case, where the newsvendor always orders $1-c$, since that quantity meets the critical-fractile condition (at that quantity, cumulative density equals $1-c$).. At $c = \frac{1}{2}$, both the average order for I_η and the average order for $I_{\eta'}$, equal $\frac{1}{2}$. So at $c = \frac{1}{2}$, the I_η graph, and the $I_{\eta'}$ graph intersect the perfect-information graph, the no-information graph, and each other.

$\boxed{\text{FIGURE 3 HERE}}$

A partial intuition for Figure 3 is provided by noting that for extremely small c, the order is very large for every signal, and the average order (over all signals) is large as well, while for c close to one, the average order is small. So it is plausible that improved information lowers average order when c is small but raises it when c is close to one. Proposition 1 supplements that intuition and covers intermediate values of c.

Illustrating Proposition 1 in a normal-distribution example. Now consider the normal-distribution example in **2.2**. Demand is distributed normally with mean μ and variance $\frac{1}{\zeta}$; that
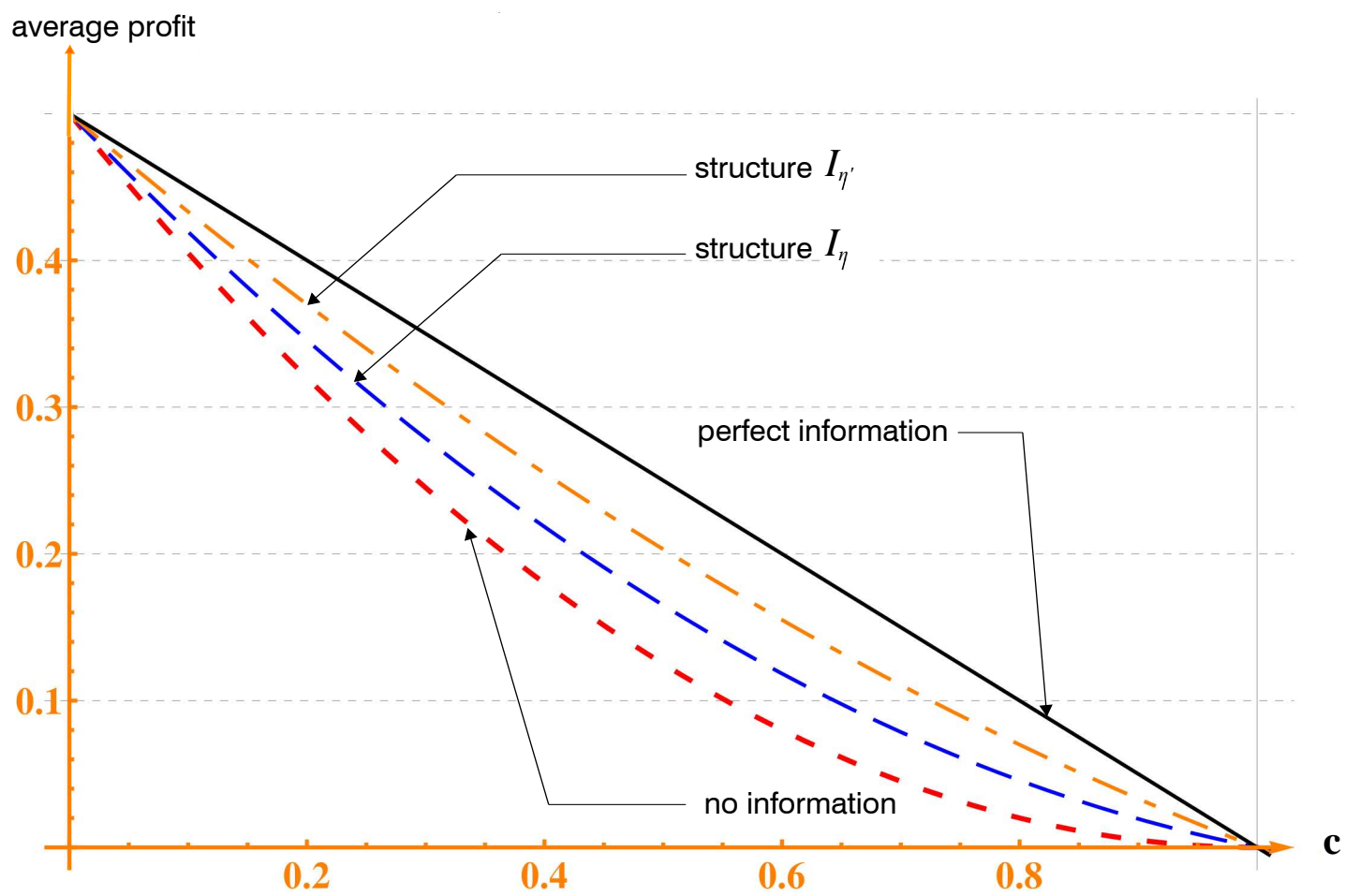
16

Figure 2

average profit: uniform example

average order

structure I_η

no information

perfect information
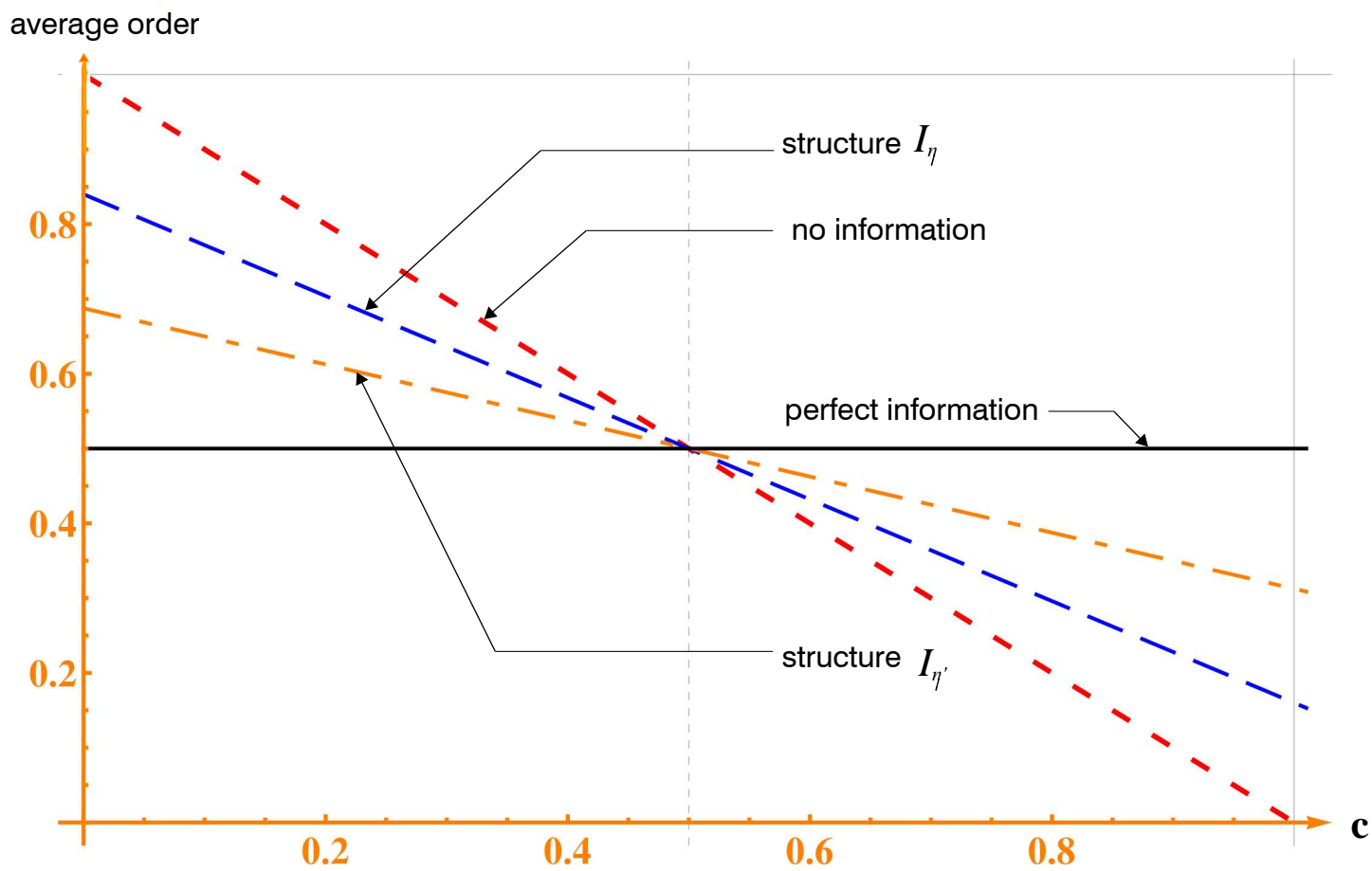
structure $I_{\eta'}$

c

Figure 3

average order: uniform example

is the newsvendor's prior. The newsvendor receives a signal from a scale/location IG whose base CDF is the standard normal, which we denote $\mathcal{N}$. When the IG has the structure I_η, then the signal $y \in Y^\eta$ tells the newsvendor that F_y^η, the posterior CDF on the demands, is the transform of the base distribution given by the location parameter $a(y, \eta)$ and the scale parameter $b(y, \eta)$. So F_y^η is normal with mean $a(y, \eta) = \dfrac{\eta y + \zeta \mu}{\eta + \zeta}$ and standard deviation $b(y, \eta) = \sqrt{\dfrac{1}{\eta + \zeta}}$. The average of the posterior mean $a(y, \eta)$, over all signals y, equals μ, the mean of the prior.

For a fixed unit cost c, consider the newsvendor's order, under the structure I_η, given the signal $y \in Y^\eta$. The order equals the mean of F_y^η plus $\mathcal{N}^{-1}(1 - c)$ times the standard deviation of F_y^η, i.e., the order equals

$$\frac{\eta y + \zeta \mu}{\eta + \zeta} + \mathcal{N}^{-1}(1 - c) \cdot \sqrt{\frac{1}{\eta + \zeta}}.$$

For the structure $I_{\eta'}$, the newsvendor's optimal order given the signal $y \in Y^{\eta'}$ is

$$\frac{\eta' y + \zeta \mu}{\eta' + \zeta} + \mathcal{N}^{-1}(1 - c) \cdot \sqrt{\frac{1}{\eta' + \zeta}}.$$

In Figure 4, the counterpart of the uniform-distribution Figure 2, we again illustrate statement [1] of the Proposition. To distinguish the structures portrayed from those of Figure 2, we label them $\tilde{I}_\eta$ and $\tilde{I}_{\eta'}$, with respective signal sets $\tilde{Y}^\eta, \tilde{Y}^{\eta'}$. For every c, the figure shows average profit (over all signals) for the two structures and also shows average profit for the perfect-information case and the no-information case. The figure is drawn for the case

$$\mu = 4, \zeta = \frac{1}{4}, \eta = \frac{7}{36}, \eta' = 2.$$

FIGURE 4 HERE

To illustrate statements [2]-[4], we continue to study this case and we calculate the average orders, $\hat{q}(\eta)$ and $\hat{q}(\eta')$. Each of the signal sets $\tilde{Y}_\eta, \tilde{Y}_{\eta'}$ is the real line. For $\tilde{I}_\eta$ and $\tilde{I}_{\eta'}$, the signals are normally distributed with mean μ and with variance $\frac{1}{\zeta} + \frac{1}{\eta}$ and $\frac{1}{\zeta} + \frac{1}{\eta'}$, respectively. The average order quantities over all signals are

$$\hat{q}(\eta) = \mu + \mathcal{N}^{-1}(1 - c) \cdot \sqrt{\frac{1}{\eta + \zeta}}, \quad \hat{q}(\eta') = \mu + \mathcal{N}^{-1}(1 - c) \cdot \sqrt{\frac{1}{\eta' + \zeta}}.$$

We compute $\hat{q}(\eta)$ and $\hat{q}(\eta')$ for every c. The graphs are shown in Figure 5, the counterpart of the uniform-distribution Figure 3. As before, the figure also depicts perfect information and no information. Under perfect information, the order equals true demand, so for every c, the average order is μ. If there is no information, then for a given c the order is the quantity meeting the critical-fractile condition for the newsvendor's prior CDF (the normal CDF with mean μ and variance $\frac{1}{\zeta}$), i.e., it is the quantity for which the mass to the left equals $1 - c$. At $c = \frac{1}{2}$, the critical-fractile quantity equals μ, and that is also the average perfect-information quantity.
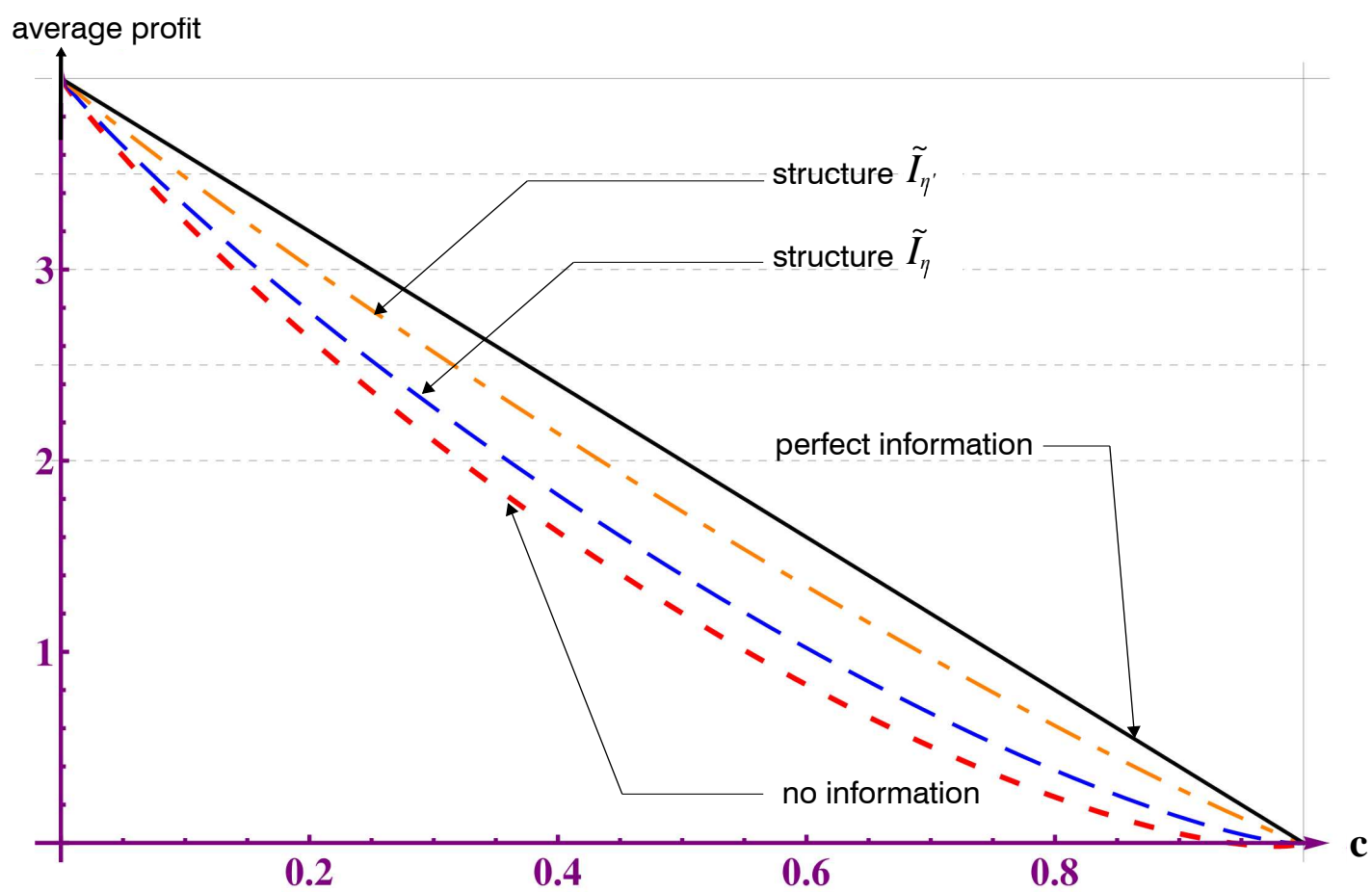
Figure 4

average profit: normal example

Figure 2 and Figure 4 share a key property: regardless of unit cost, the higher effort structure yields higher average profit (over all signals). Figures 3 and 5 share a second key property: there is a critical value of c such that the high-effort average order (over all signals) is less than the low-effort average order for costs below the critical value but is greater for costs above it. Both properties would also be exhibited in a pair of figures that we omit here. They concern the triangular example in Section 2.3. On the other hand, in the three-demand example of the Introduction, where (as we saw) the scale/location property fails, both of the key properties are violated. For average order we have Figure 1, which stands in sharp contrast to Figures 3 and 5. We obtain Figures 2 - 5 by direct computation. For the uniform-distribution figures (Figure 2 and Figure 3), we use the formulae given in (2) and (3). Another direct computation yields the omitted figures concerning the triangular case. If we had *only* the figures, we would be puzzled by their sharing of the key properties. The puzzle is resolved by Proposition 1, where a unifying argument covers the uniform-distribution example, the normal-distribution example, the triangular example, and others. The common thread is that every newsvendor posterior can be interpreted as the scale/location transform of a fixed base distribution. As we have seen, that fact turns out to articulate well with the key property of the newsvendor's payoff function — the function is shift-invariant and satisfies a homogeneity condition. The three-demand example of the Introduction lacks the scale/location property and it lacks the key properties of Proposition 1 as well.

The next proposition concerns the classic newsvendor's average *ex post* sales. This time there is no critical value of c. Consider a switch in IG structure which is useful to the newsvendor and hence (as statement [1] of Proposition 1 showed), reduces the average scale parameter. Whatever c may be, the switch increases the average, over all signals, of *ex post* sales.

In stating the proposition, we let $\hat{S}(\eta)$ denote the average *ex post* quantity, over all the signals in Y^η, when the newsvendor responds to each signal with an optimal order. Thus

$$\hat{S}(\eta) \equiv E_{y \in Y^\eta} \Big(E_{T_{a(y,\eta),b(y,\eta)}} \min \big(q_{a(y,\eta),b(y,\eta)}, D \big) \Big).$$

Proposition 2

Let the assumptions of Proposition 1 hold. Then

$$U(\eta') > U(\eta) \text{ if and only if } \hat{S}(\eta') > \hat{S}(\eta).$$

Illustrating Proposition 2 in a uniform-distribution example. Another straightforward argument shows that if $[0, 1]$ is partitioned into n subintervals with widths $p_1, \ldots, p_n$, then

(4) $$\text{average } ex \ post \text{ sales over all subintervals } = \frac{1}{2} - \frac{1}{2}c^2 \cdot \sum_{i=1}^{n} (p_i)^2.$$

18

average order

structure $\tilde{I}_\eta$

no information

perfect information

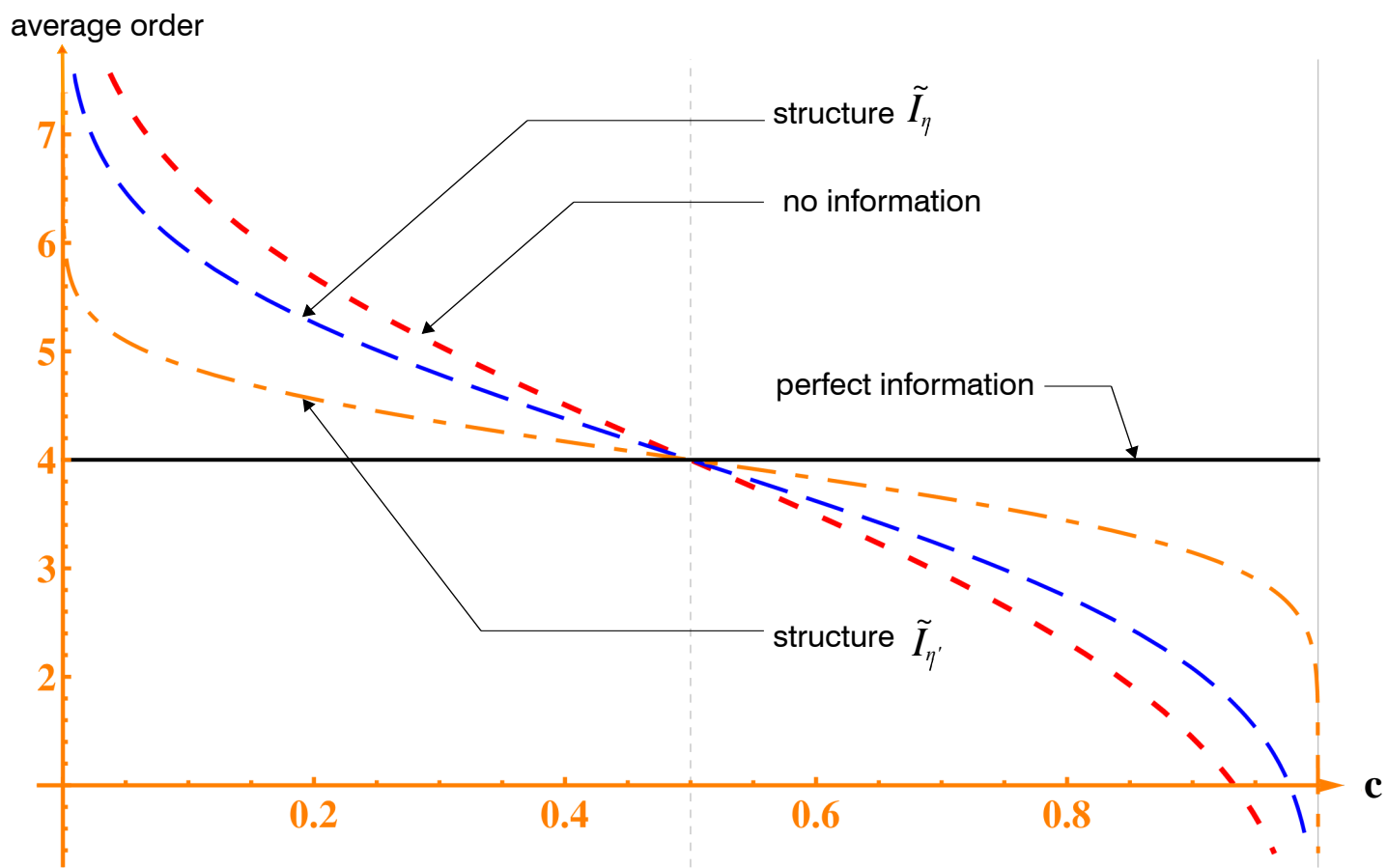structure $\tilde{I}_{\eta'}$

c

Figure 5

average order: normal example

Using (4), we calculate average *ex post* sales for the same two structures I_η and $I_{\eta'}$ that are used in Figures 2 and 3. Figure 6 then graphs average *ex post* sales as a function of c for each structure. Once again, there are also graphs for perfect information and for no information.

FIGURE 6 HERE

Illustrating Proposition 2 in a normal-distribution example. A somewhat cumbersome calculation (using the formula for the mean of a left-tail truncated normal distribution) shows that for our normal-distribution example

$$(5) \qquad \text{average } \textit{ex post} \text{ sales } = \mu + \sqrt{\frac{1}{\eta + \zeta}} \cdot \left[\int \min\left(x, \mathcal{N}^{-1}(1 - c)\right) \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \right].$$

Note that

$$\int \min\left(x, \mathcal{N}^{-1}(1 - c)\right) \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx < \int x \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0.$$

So average *ex post* sales rise when η rises. That is seen in Figure 7, where we again use $\mu = 4, \zeta = \frac{1}{4}, \eta = \frac{7}{36}, \eta' = 2$. Again, the figure has two additional graphs, for the perfect-information case and the no-information case.

FIGURE 7 HERE

As before, we can study *ex post* sales directly for the uniform-distribution example (using (4)) and for the normal-distribution example (using (5)). But Proposition 2 again provides a unifying argument that covers both examples and explains why Figure 6 and Figure 7 share a key property that is lacking in the three-demand example of the Introduction. In that example, as we saw, the IG does not perform a scale/location transform. Higher IG effort yields higher average profit, but for some values of c, average *ex post* sales *fall* when effort increases. Proposition 2 tells us that this cannot happen when the IG is a scale/location transformer. Then higher effort yields higher average profit and for every c, average *ex post* sales rise when effort increases.

5. A generalization: a newsvendor who is "regret-averse"

5.1 A class of quantity-choosers that includes a regret-averse newsvendor.

Consider next a new version of the newsvendor. To introduce the new version, we first recall that the *classic* newsvendor's task can be stated in an alternative way: he seeks to minimize

$$c \cdot (q - D)^+ + (1 - c) \cdot (D - q)^+,$$

where J^+ is the standard shorthand for $\max(0, J)$. It must be the case that one of the two summed terms equals zero. The first term is the unsold-inventory (overage) penalty and the second term is the unfulfilled-demand (underage) penalty. We can rewrite the whole expression as

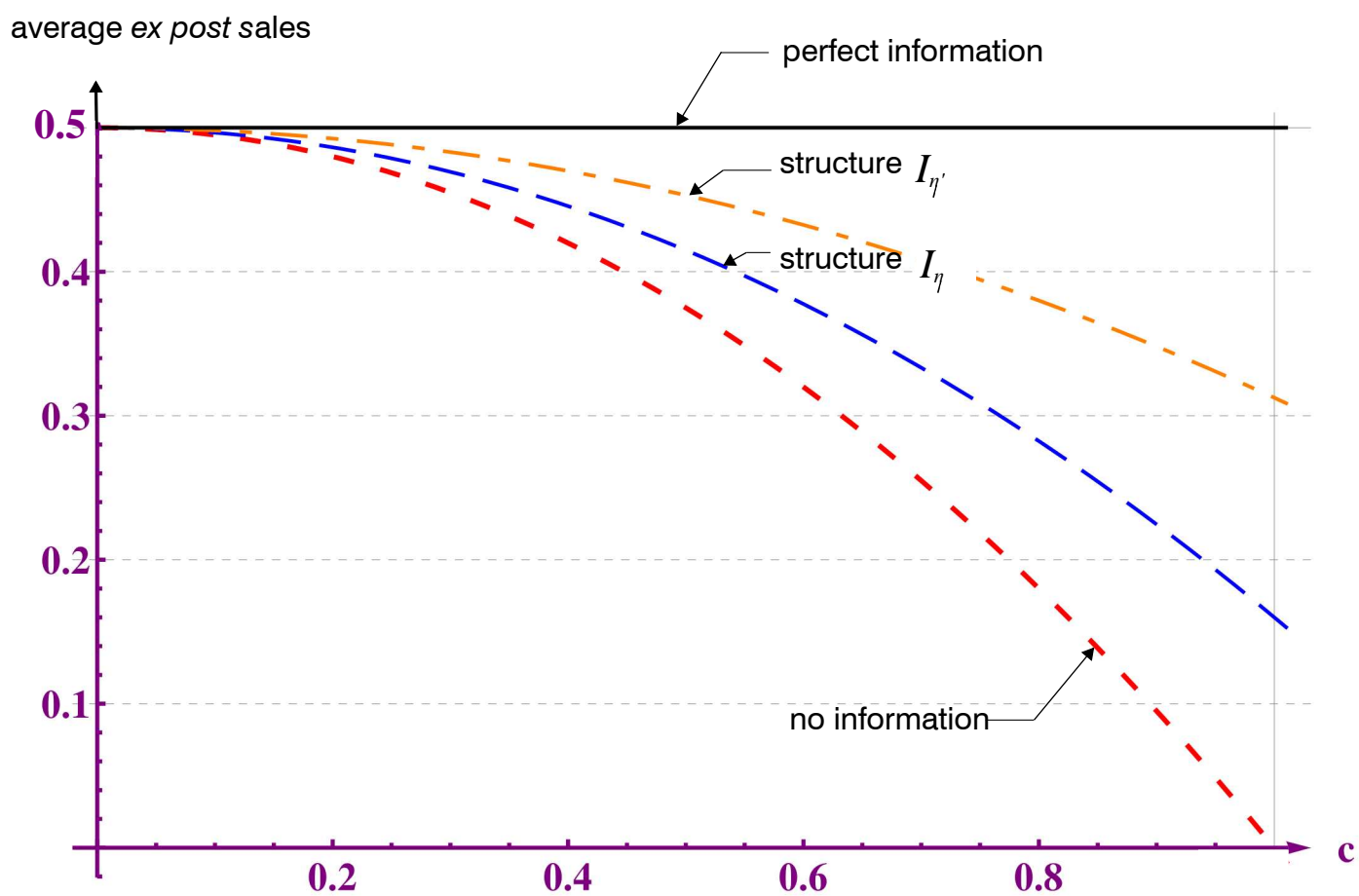$$c \cdot [(q - D)^+ + \frac{1 - c}{c} \cdot (D - q)^+].$$

19

average *ex post* sales

perfect information

structure $I_{\eta'}$

structure I_{η}

no information

0.5

0.4

0.3

0.2

0.1

0.2   0.4   0.6   0.8   c

Figure 6

average *ex post* sales: uniform example

average *ex post* sales

perfect information

structure $\widetilde{I}_{\eta'}$

structure $\widetilde{I}_{\eta}$

no information
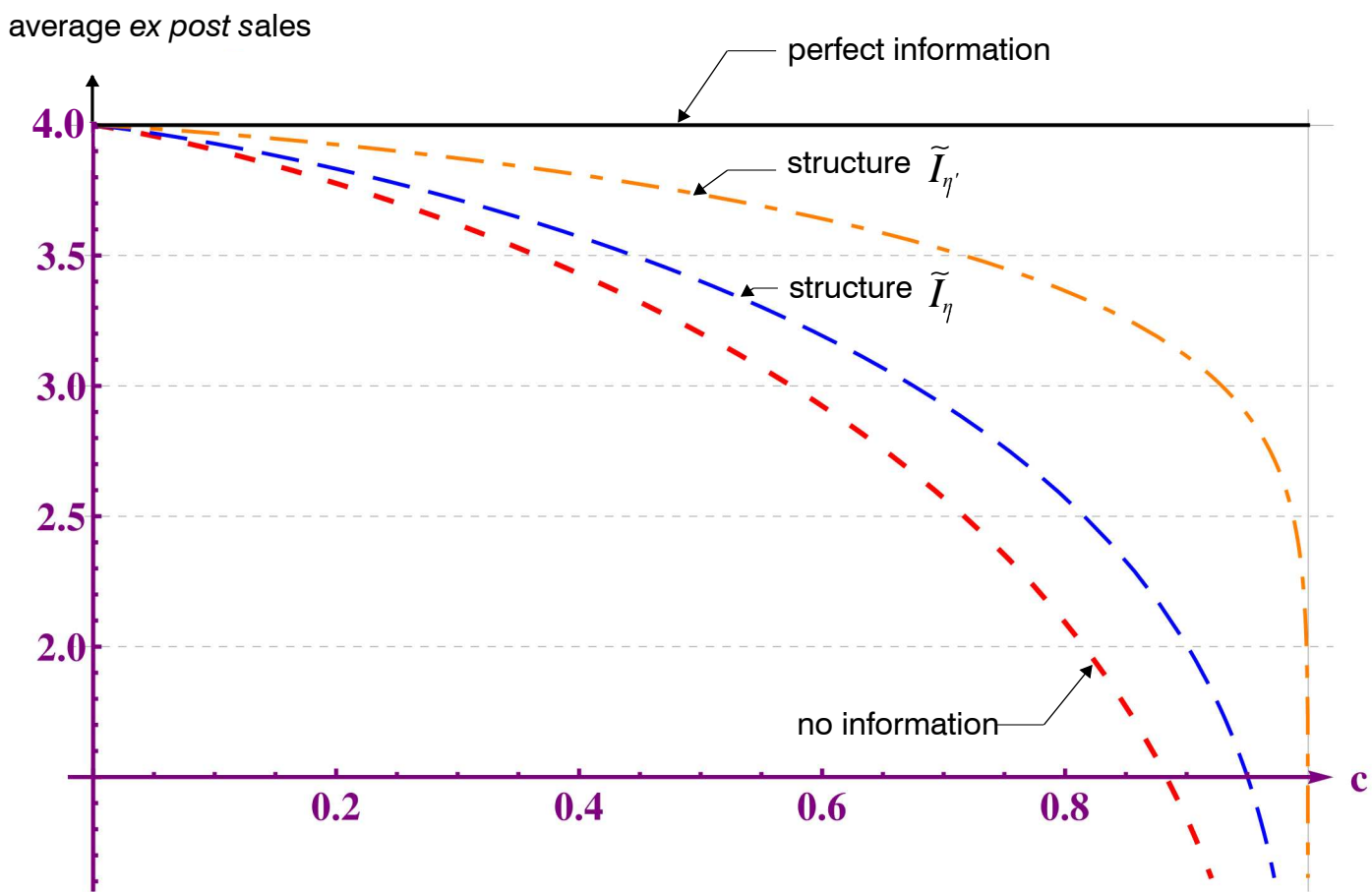
4.0

3.5

3.0

2.5

2.0

0.2   0.4   0.6   0.8

c

Figure 7

average *ex post* sales: normal example

Thus — since $c > 0$ — the classic newsvendor seeks to maximize the following payoff function:

(6)
$$u(q, \theta) = \begin{cases} -|q - \theta|^r & \text{for } q \geq \theta \\ -K \cdot |q - \theta|^r & \text{for } q < \theta, \end{cases}$$

where $\theta = D$, $K = \left(\frac{1-c}{c}\right)^r$ and $r = 1$.

Now suppose that when the newsvendor's quantity deviates from the quantity that turns out to be ideal — namely the quantity that equals realized demand — then the penalty suffered is no longer linear in the absolute size of the deviation. Instead the penalty is given in the above payoff function u, where θ again equals D, K again equals $\left(\frac{1-c}{c}\right)^r$, but r is now *greater than one*. Such a newsvendor is our *generalized* newsvendor.

It will be useful to let the state variable in the function u continue to be the above general variable θ and to study a quantity-chooser whose payoff function u is defined by (6), where K is an arbitrary positive number. The state θ chosen by Nature is the quantity that turns out to be ideal, and u describes the penalties suffered by a quantity-chooser when the quantity he chooses departs from the ideal. Our main interest is the case where $\theta = D$, the quantity-chooser is our generalized newsvendor, and $K = \left(\frac{1-c}{c}\right)^r$. But the Proposition which follows is stated for a general quantity-chooser.

We note that the generalized newsvendor described by the above u is *regret-averse*. Consider regret in a general setting. Given a payoff function, the regret associated with a state and an action taken before that state becomes known is the amount by which the payoff actually earned falls short of what could have been earned had the state been known before the action had to be chosen. An early and widely cited discussion of regret appears in Bell (1983). In our setting, continue to assume that price is one and cost is c with $0 < c < 1$. Then regret when the order quantity is q and demand turns out to be D is

$$\rho(q, D) = c \cdot \max(q - D, 0) + (1 - c) \cdot \max(D - q, 0).$$

Note that either the first or the second of the two summed terms must equal zero. The first term of ρ is the regret experienced when demand becomes known and turns out to be less than or equal to the order quantity q. The second term is the regret experienced when demand turns out to be greater than or equal to q. Regret is zero when demand turns out to match q. But since one of the two terms in ρ has to equal zero, we have, for any $r > 1$ and for the payoff function u in (+) (with $K = \left(\frac{1-c}{c}\right)^r$):

$$\begin{aligned} [\rho(q, D)]^r &= c^r \cdot [(q - D)^+]^r + (1 - c)^r \cdot [(D - q)^+]^r \\ &= -c^r \cdot u(q, D). \end{aligned}$$

Thus maximizing the expected value of u (and hence the expected value of $c^r \cdot u$) is equivalent to minimizing the expected value of the rth power of regret. A decision-maker who minimizes the expected value of regret raised to the power r, where $r > 1$, is said to be *regret-averse*.

Various departures from the classic risk-neutral newsvendor have been explored. The expected-utility-maximizing risk-averse newsvendor is discussed in Agrawal and Seshadri (2000); Eekhoudt,

Gollier, and Schlesinger (1995); Lau (1980); and Choi asnd Ruszczynski (2007). The newsvendor who minimizes maximum regret is discussed in Perakis and Roels (2008). Thus far, however, the regret-averse newsvendor has not appeared on the scene. Papers by experimentalists, like Schweitzer and Cachon (2000) and Ho, Lim, and Cui (2010), report that some subjects find avoiding unsold stock to be more important than avoiding unfulfilled demand. The regret-averse newsvendor would seem to be a promising model in studying the way the strength of this bias changes when there is a change in c or in information about demand. Moreover a recent paper by Hayashi (2008) provides axioms that justify regret aversion.

Note next that the payoff function in (6) satisfies a condition which generalizes the shift invariance and homogeneity condition in (1) above. A payoff function $u(q, \theta)$ satisfies the generalized condition if:

$$(1') \quad \begin{cases} \text{There exists a function } f \text{ such that for any } \alpha \text{ and any } \beta > 0, \\[2mm] (\dagger) \quad f(\beta) > 0 \text{ and } u(q, \theta) = f(\beta) \cdot u\left(\frac{q - \alpha}{\beta}, \frac{\theta - \alpha}{\beta}\right) \text{ for all } q, \theta. \end{cases}$$

For the case where f is the identity, condition $(1')$ becomes condition (1), where the "γ" of condition (1) is zero.

Our general payoff function u in (6) obeys condition $(1')$ with $f(\beta) = \beta^r$. To see this, first consider pairs (q, θ) such that $q \geq \theta$. Then for $\beta > 0$ we have $\frac{q-a}{\beta} \geq \frac{\theta-a}{\beta}$ and hence

$$u(q, \theta) = -|q - \theta|^r = \beta^r \cdot \left|\frac{q - a}{\beta} - \frac{\theta - a}{\beta}\right|^r = \beta^r \cdot u\left(\frac{q - a}{\beta}, \frac{\theta - a}{\beta}\right).$$

For pairs (q, θ) such that $q < \theta$, we have $\frac{q-a}{\beta} < \frac{\theta-a}{\beta}$ and hence

$$u(q, \theta) = -K \cdot |q - \theta|^r = \beta^r \cdot \left[-K \cdot \left|\frac{q - a}{\beta} - \frac{\theta - a\cdot}{\beta}\right|^r\right] = \beta^r \cdot u\left(\frac{q - a}{\beta}, \frac{\theta - a}{\beta}\right).$$

We now have a general proposition about a quantity-chooser. In particular, the quantity-chooser may be our generalized newsvendor. Then the proposition becomes a counterpart of our classic-newsvendor Proposition 1. Its proof uses an extended version of Lemma A and of Lemma B as well as the fact that the payoff function satisfies $(1')$. The proposition again considers a switch from one IG structure to another. We require, however, that when the switch occurs, the average (over all signals) of the scale parameter b moves in the same direction as the average of b^r. If that is so, then we again find that the new structure is more useful for the quantity-chooser than the old one if and only if the average scale parameter drops after the switch.

In addition, the proposition finds that there is a critical value of the penalty parameter K. On one side of that value, a useful switch increases the quantity-chooser's average *ex ante* quantity. On the other side, such a switch decreases the average quantity. Finally we find that if the base CDF B has a density function that is symmetric around zero (or if $r = 2$) then the critical

value of K is one. If the quantity-chooser is our regret-averse newsvendor, then $K = \left(\frac{1-c}{c}\right)^r$. Hence $K = 1$ is equivalent to $c = \frac{1}{2}$. So for the symmetric case we have generalized-newsvendor counterparts of the classic-newsvendor statements [2] and [3] of Proposition 1.

In stating the proposition we again use the symbols $\hat{q}(\ \cdot\ ), U(\ \cdot\ ), V(\ \cdot\ )$. In defining those symbols (just before Lemma B) we used the quantity $q_{a(y,\eta),b(y,\eta)}$, the smallest maximizer of $E_{T_{a(y,\eta),b(y,\eta)}} u(q, D)$. Our new definitions of $\hat{q}(\ \cdot\ ), U(\ \cdot\ ), V(\ \cdot\ )$ again use the quantity $q_{a(y,\eta),b(y,\eta)}$, in exactly the same way as before, but that quantity is now the smallest maximizer of $E_{T_{a(y,\eta),b(y,\eta)}} u(q, \theta)$, where $u(q, \theta)$ is the payoff function in (+). We also use a new symbol, namely $V_r(\eta)$. That denotes the average, over all the signals in Y^η, of the rth power of the scale parameter $b(y, \eta)$.

Proposition 3

Consider a quantity-chooser whose payoff function is the function defined in (6), namely

$$(6) \qquad\qquad u(q, \theta) = \begin{cases} -|q - \theta|^r & \text{for } q \geq \theta \\ -K \cdot |q - \theta|^r & \text{for } q < \theta, \end{cases}$$

where θ is a state variable, $r > 1$ and $K > 0$. Consider an Information-gatherer who obtains scale/location transforms of a base CDF B which has mean zero, is not a one-point distribution, and has a density that is positive at every point of its support. Every signal y in the IG's signal set Y^η defines the posterior CDF used by the quantity-chooser, namely the transform $F_y^\eta = T_{a(y,\eta),b(y,\eta)}$. Let $I_\eta = \langle Y^\eta, \{F_y^\eta\}_{y \in Y^\eta}\rangle$ and $I_{\eta'} = \langle Y^{\eta'}, \{F_y^{\eta'}\}_{y \in Y^{\eta'}}\rangle$ be two of the IG's structures. Suppose that

$(*)$ when the IG switches from I_η to $I_{\eta'}$, then V and V_r move in the same direction.

Then

[1] $V(\eta') < V(\eta)$ if and only if $U(\eta') > U(\eta)$.

Moreover, there exists K^* such that

[2] If $K > K^*$, then
$$U(\eta') > U(\eta) \text{ if and only if } \hat{q}(\eta') > \hat{q}(\eta).$$

[3] If $K < K^*$, then
$$U(\eta') > U(\eta) \text{ if and only if } \hat{q}(\eta') < \hat{q}(\eta).$$

[4] If $r = 2$, then statements [2] and [3] hold for $K^* = 1$.

[5] If B is symmetric around zero — i.e., for any z we have $B(-z) = 1 - B(z)$ — then statements [2] and [3] hold for $K^* = 1$.

Note that if the quantity-chooser is our generalized newsvendor — so $K = \left(\frac{1-c}{c}\right)^r$ — and if $r \neq 2$ and B is *not* symmetric around zero, then [2], [3] are statements about a critical value of c and that critical value is *not* $\frac{1}{2}$. The inequalities $K > K^*, K < K^*$ are equivalent, respectively, to the inequalities $c < \frac{1}{(K^*)^{1/r}+1}, c > \frac{1}{(K^*)^{1/r}+1}$.

To interpret condition $(*)$, which concerns the IG, first consider the IG in the normal-distribution example in 2.2 above. Recall that $b(y, \eta) = \sqrt{\frac{1}{\eta+\zeta}}$. The signal y does not enter that expression, so condition $(*)$ is automatically met. Next consider the IG in the uniform-distribution example in 2.1. The set of possible demands is $[0, 1]$. Suppose that the signal set Y^η partitions $[0, 1]$ into η *equal* subintervals. We may let Y^η be the first η positive integers. The signal $y = i$ tells the newsvendor that the true D is uniformly distributed on $\left[a(i, \eta) - \frac{b(i, \eta)}{2}, a(i, \eta) + \frac{b(i, \eta)}{2}\right]$, where $a(i, \eta) = \frac{2i-1}{2\eta}$ and $b(i, \eta) = \frac{1}{\eta}$. Thus

$$V(\eta) = E_{i \in Y^\eta} \, b(i, \eta) = \frac{1}{\eta}; \quad V_r(\eta) = E_{i \in Y^\eta} \, (b(i, \eta))^r = \left(\frac{1}{\eta}\right)^r.$$

So when the index η changes, V and V_r indeed move in the same direction.

Suppose, on the other hand, that the IG does not partition $\mathcal{D} = [0, 1]$ into equal subintervals. Rather there is a fixed integer $m > 0$ such that for every η the signal set Y^η defines a partitioning of $[0, 1]$ having m subintervals whose widths need not be equal. The m-subinterval partitioning changes when η changes. For a given η, $V(\eta)$ is the average of the m subinterval widths, while $V_r(\eta)$ is the average of the m widths, each raised to the power r. Without further restrictions we cannot claim that V and V_r move in the same direction when η changes.

It is easily verified, however, that we *can* make that claim (for any $r > 1$) if $m = 2$, i.e., every signal set Y^η partitions $[0, 1]$ into just *two* subintervals.

For any fixed m, moreover, we can claim that V and V_r move in the same direction if the IG's information structures I_η obey the following condition:

> Suppose $\eta' > \eta$; $w' = (w'_1, \ldots, w'_m)$ is the vector of subinterval widths for $Y^{\eta'}$; and $w = (w_1, \ldots, w_m)$ is the vector of subinterval widths for Y^η. Then w' *majorizes* w — i.e., the largest component of w' is not lower than the largest component of w, the sum of the two largest is not lower, the sum of the three largest is not lower, and so on.

It is shown in the Appendix that this claim is correct.

5.2 An *ex post* proposition for the generalized newsvendor.

Let our quantity-chooser again be our generalized regret-averse newsvendor. We can repeat a key step in the proof of Proposition 3 to obtain an analog of our Proposition 2, which concerned *ex post* sales for a classic newsvendor. Suppose the IG moves from one information structure to another and the second is more useful to the newsvendor than the first. (That is equivalent, as [1] of Proposition 3 told us, to the statement that the IG's average scale parameter drops). Then

23

we find, just as in Proposition 2, that for *all* values of the cost c (with $0 < c < 1$), we have an increase in the average (over all the IG's signals) of the generalized newsvendor's *ex post* sales.

To state this final proposition, we need to reinterpret the symbol

$$\hat{S}(\eta) \equiv E_{y \in Y^\eta} \left(E_{T_{a(y,\eta),b(y,\eta)}} \min \left(q_{a(y,\eta),b(y,\eta)}, D \right) \right).$$

The quantity $q_{a(y,\eta),b(y,\eta)}$ is now the smallest maximizer of $E_{T_{a(y,\eta),b(y,\eta)}} u(q, D)$, where $u(q, D)$ is the payoff function in (6) (with $\theta = D$).

Proposition 4

Consider again the generalized newsvendor whose payoff function given in (6), i.e., it is

$$u(q, D) = \begin{cases} -|q - D|^r & \text{for } q \geq D \\ -K \cdot |q - D|^r & \text{for } q < D, \end{cases}$$

where $K = \left(\frac{1-c}{c} \right)^r$ and $r > 1$. Consider an Information-gatherer who obtains scale/location transforms of a base CDF B which has mean zero and a density that is positive at every point of its support. Every signal y in the IG's signal set Y^η defines the posterior used by the newsvendor, namely $F_y^\eta = T_{a(y,\eta),b(y,\eta)}$. Let $I_\eta = \langle Y^\eta, \{F_y^\eta\}_{y \in Y^\eta} \rangle$ and $I_{\eta'} = \langle Y^{\eta'}, \{F_y^{\eta'}\}_{y \in Y^{\eta'}} \rangle$ be two of the IG's structures. Then

$$U(\eta') > U(\eta) \text{ if and only if } \hat{S}(\eta') > \hat{S}(\eta).$$

6. Concluding remarks.

We have considered a newsvendor who obtains the services of an Information-gatherer. The IG expends effort and sends to the newsvendor a signal that provides information about the unknown true demand. A joint distribution on the possible signal/demand pairs is known to the IG and the newsvendor. The IG's signal defines a posterior and the newsvendor uses that posterior to choose an order quantity. We studied the effect of more IG effort on the average *ex ante* order and the average *ex post* sales.

We developed a model of the IG for which (1) there is a critical value of the classic newsvendor's unit cost, where on one side more effort leads to a higher average *ex ante* order and on the other side to a lower one; and (2) regardless of cost, more IG effort always leads to higher average *ex post* sales. Moreover, (1) and (2) continue to hold when we replace the classic newsvendor with a generalized (regret-averse) newsvendor. In our model each of the posteriors which the newsvendor obtains from the IG is a scale/location transform of a fixed base distribution, and more effort means a drop in the average scale parameter.

More remains to be learned about scale/location IGs. In the Introduction we considered a three-demand example where the posteriors implied by the IG's signals were not transforms of

24

a common base distribution. But, as we saw in the very simple finite example in 2.3, we can also construct finite examples where the scale/location property holds. The effect of more IG effort on the newsvendor can be studied in more general finite cases. When we turn to IGs who are *not* scale/location transformers, there is even more to be learned. It is reasonable to confine attention to models in which higher IG effort is useful to the newsvendor. Even then, there are IG models for which statements (1) and (2) no longer hold.

Rapid improvements in the technology of information-gathering in general, and demand forecasting in particular, provide a stimulus for further research. As surveying of consumers gets cheaper, it becomes particularly interesting to study the case where the IG is the sampler briefly discussed in the Introduction He samples the demanders and determines how much each demander in the sample is willing to pay for a given amount. As sample size grows, do average order and average sales rise or fall?

Empirical studies of real information-gatherers and their efforts remain quite scarce. That may be because theory has been scarce as well. Without theory, empiricists may not know what hypotheses to test and what datasets to assemble. If theorists show that for a particular model of an IG more effort induces information users to move their choices in a certain direction (induces newsvendors to increase their orders, for example), then that tells the empiricist that it would be useful to see whether real IGs fit that particular model and whether real users' choices indeed move in the direction that the theory predicts. In another paper (Marschak, Shanthikumar, and Zhou (2013)) we study a variety of IGs who are not scale/location transformers and a variety of "Producers", who use the IG's signals and are not newsvendors. For each IG/Producer pair we ask the same question that we have asked in the present paper: does more IG effort raise or lower the average quantity chosen by the Producer?

It is not only the model of the IG that can be varied but the model of the newsvendor as well. For example, the newsvendor might choose both price and quantity, as in Agrawal and Seshadri (2000). There is a randomly shifting demand curve which determines realized demand for a given price, and an IG can be asked to study the shifts. If the newsvendor's chosen quantity is less than realized demand, then he can buy more at a fixed "emergency" price; if it is more, he receives a salvage value for the unsold quantity. In another variation, we can continue to let quantity be the only choice variable, with selling priced fixed and quantity demanded uncertain, but we can try to generalize to a newsvendor who is risk-averse rather than regret-averse.

We have not asked how much IG effort the newsvendor ought to buy if he has to pay for it. We have also avoided incentive-oriented models where the IG chooses how hard to work, bears the cost, and receives a reward from the newsvendor. In our view those hard questions are important, but it is appropriate to defer them until we better understand, for a variety of IG models, the effect of more effort on the quantities we have studied.

APPENDIX

1. Proof of Lemma A

First note that for any function g, a change-of-variables argument yields

(i) $$E_B \; g(a + br) = E_{T_{ab}} \; g(r).$$

Now fix q, let D play the role of "r", and let $u(q, \cdot)$ play the role of "$g(\cdot)$". Using (i), we obtain

$$E_{T_{ab}} \; u(q, D) = E_B \; u(q, a + bD).$$

But since u has the property (1) (defined in Section 3), we have

$$E_B \; u(q, a + bD) = \gamma a + b \cdot E_B \; u\left(\frac{q - a}{b}, D\right),$$

where $\gamma = 1 - c$. Thus

(ii) $$E_{T_{ab}} \; u(q, \theta) = \gamma a + b \cdot E_B \; u\left(\frac{q - a}{b}, D\right).$$

Since $b > 0$, (ii) implies that a quantity q which maximizes $E_B \; u\left(\frac{q-a}{b}, D\right)$ also maximizes $E_{T_{ab}} \; u(q, D)$. Hence

(iii) $$u_{T_{ab}} = \max_q E_{T_{ab}} \; u(q, D) = \gamma a + b \cdot \max_q E_B \; u\left(\frac{q - a}{b}, D\right).$$

But for fixed (a, b),

(iv)   the set of possible values of $\frac{q-a}{b}$ equals the set of possible values of q.

(Both sets consist of all the reals). Thus we have

$$\max_q E_B \; u\left(\frac{q - a}{b}, D\right) = \max_{(q-a)/b} E_B \; u\left(\frac{q - a}{b}, D\right) = \max_q E_B \; u(q, D) = u_B.$$

So, using (iii),

$$u_{T_{ab}} = \gamma a + b \cdot u_B,$$

i.e., (A2) holds.

Next note that since q_{ab} is a value of the variable q which maximizes the term on the left of the equality in (ii), the quantity $\frac{q_{ab} - a}{b}$ maximizes the entire term on the right of the equality in (ii) and therefore (since $b > 0$) it maximizes $E_B \; u\left(\frac{q-a}{b}, \theta\right)$ as well. So we obtain, using (iv),

$$E_B \; u\left(\frac{q_{ab} - a}{b}, D\right) \geq E_B \; u(q, D) \text{ for all } q \in \mathbb{R}.$$

Since q_B is the smallest maximizer of $E_B \; u(q, D)$, we have

(v) $$q_B \leq \frac{q_{ab} - a}{b}.$$

Now apply (ii) to obtain

(vi) $$E_{T_{ab}} \, u(a + bq_B, D) = \gamma a + b \cdot E_B \, u(q_T, D).$$

Since $E_B \, u(q_B, D) \geq E_B \, u(q, D)$ for all $q \in \mathbb{R}$, (vi) implies that

$$E_{T_{ab}} \, u(a + bq_B, D) \geq E_{T_{ab}} \, u(q, D) \text{ for all } q \in \mathbb{R}.$$

Since q_{ab} is the smallest maximizer of $E_{T_{ab}} \, u(q, D)$, we have

$$q_{ab} \leq a + bq_B.$$

Since (v) implies $q_{ab} \geq a + bq_B$, we conclude that

$$q_{ab} = a + bq_B,$$

i.e., (A1) holds. □

2. Proof of Lemma B.

First recall that for a given signal y, the CDF $T_{a(y,\eta), b(y,\eta)}$ is the (posterior) CDF of the random variable D, and that D is a transform of the base random variable z. Specifically, for a given signal y, we have

$$D = a(y, \eta) + b(y, \eta) \cdot z,$$

where $b(y, \eta) > 0$. Since the mean of the base random variable is zero, the transformed random variable D has (posterior) mean

(i) $$E_{T_{a(y,\eta), b(y,\eta)}} \, D = a(y, \eta).$$

Next we use the following general fact, which holds for any IG:

(#)
$\left\{ \begin{array}{l} \text{The expectation, over all the Information-gatherer's signals, of the pos-} \\ \text{terior mean given the signal equals the prior mean, i.e., if } Y \text{ is the set} \\ \text{of possible signals, we have} \\ \\ \qquad\qquad E_{y \in Y}\left(E_{F_y} D \right) = \bar{D}, \\ \\ \text{where } \bar{D} \text{ is the mean of the prior CDF and } F_y \text{ is the posterior given} \\ \text{the signal } y. \end{array} \right.$

Applying (#), we obtain

(ii) $$E_{y \in Y^\eta}\left(E_{T_{a(y,\eta), b(y,\eta)}} \, D \right) = \bar{D}.$$

Using (i) and (ii), we obtain

(iii) $$E_{y \in Y^\eta} \, a(y, \eta) = \bar{D}.$$

27

(A1) of Lemma 1 implies, for every (y, η),

$$q_{a(y,\eta),b(y,\eta)} = a(y,\eta) + b(y,\eta) \cdot q_B.$$

Hence, taking expectations over the signals $y \in Y^\eta$, and using (iii), we obtain

(iv) $$\hat{q}(\eta) = \bar{D} + q_B \cdot V(\eta).$$

Moreover (i), (ii), and (A2) of Lemma 1 imply

(v) $$U(\eta) = \gamma \bar{D} + u_B \cdot V(\eta).$$

Statement (iv) implies (B1) and statement (v) implies (B2). $\quad \square$

3. Proving Proposition 1.

In proving Proposition 1, we will need the following Lemma.

Lemma C

Suppose the IG is a scale/location transformer with a family of structures $I_\eta = \langle Y^\eta, \{F_y^\eta\}_{y \in Y^\eta} \rangle$. Then the following statements hold.

(C1)   If $u_B < 0$ then $V(\eta') < V(\eta)$ if and only if $U(\eta') > U(\eta)$.

(C2)   If $q_B < 0$, then $V(\eta') < V(\eta)$ if and only if $\hat{q}(\eta') > \hat{q}(\eta)$.

(C3)   If $q_B > 0$, then $V(\eta') < V(\eta)$ if and only if $\hat{q}(\eta') < \hat{q}(\eta)$.

(C4)   If $u_B q_B > 0$ then $U(\eta') > U(\eta)$ if and only if $\hat{q}(\eta') > \hat{q}(\eta)$.

(C5)   If $u_B q_B < 0$ then $U(\eta') > U(\eta)$ if and only if $\hat{q}(\eta') < \hat{q}(\eta)$.

(C6)   If $u_B = 0$ and $q_B \neq 0$, then $U(\eta)$ is the same for all η.

(C7)   If $u_B \neq 0$ and $q_B = 0$, then $\hat{q}(\eta)$ is the same for all η.

(C8)   If $u_B = 0$ and $q_B = 0$, then $\hat{q}(\eta)$ is the same for all η and so is $U(\eta)$.

Proof:

(C1) follows from (B2) of Lemma B.
(C2) and (C3) follow from (B1) of Lemma B.
(C4) - (C8) follow from (B1) and (B2) of Lemma B.   □

Proof of Proposition 1

By assumption the base CDF B has an everywhere positive density function, which we shall denote g, That implies that $B^{-1}(1-c)$ is the unique maximizer of $E_B\left[\min(q,\theta) - cq\right]$.

<u>Step 1</u>

In this Step we show that $u_B < 0$.

Since q_B maximizes $E_B\left(\min(q,D) - cq\right)$ it satisfies $B(q_B) = 1 - c$. We have

$$
\begin{aligned}
u_B &= \int_{-\infty}^{\infty} [\min(q_B, z) - cq_B] g(z) dz \\
&= \int_{-\infty}^{q_B} zg(z) dz + \int_{q_B}^{\infty} q_B g(z) dz - cq_B \\
&= \int_{-\infty}^{q_B} zg(z) dz + q_B \cdot (1 - B(q_B)) - cq_B \\
&= \int_{-\infty}^{q_B} zg(z) dz.
\end{aligned}
$$

(To obtain the final equality, we use the fact that $B(q_B) = 1 - c$). If $q_B \leq 0$, then at every z in $[-\infty, q_B]$ we have $zg(z) \leq 0$. The set $\{z \ : \ zg(z) < 0\}$ has positive probability.

So if $q_B \leq 0$, then
$$
u_B = \int_{-\infty}^{q_B} zg(z) dz < \int_{-\infty}^{q_B} 0 \ dz = 0.
$$
Suppose, on the other hand, that $q_B > 0$. Then
$$
u_B = \int_{-\infty}^{q_B} zg(z) dz = \int_{-\infty}^{\infty} zg(z) dz - \int_{q_B}^{\infty} zg(z) dz.
$$

The second integral equals zero since the CDF B has zero mean. The third integral is positive since for every z in $[q_T, \infty]$ the integrand is positive.

Thus whatever the sign of q_B, we have $u_B < 0$.

<u>Step 2</u>

Since $u_B < 0$, we immediately obtain statement **[1]** of the Proposition from (C1) of Lemma C.

<u>Step 3</u>

In this Step we show **[2]**, **[3]**, **[4]**.

If $1 - c < B(0)$ then $q_B < 0$ (since the CDF B has mean zero and q_B satisfies $B(q_B) = 1 - c$). Since $u_B < 0$, we have $u_B \cdot q_B > 0$. Using (C4) of Lemma C we obtain statement [2].

If $1 - c > B(0)$ then $q_B > 0$. Since $u_B < 0$, we have $u_B \cdot q_B < 0$. Using (C5) of Lemma C we obtain statement [3].

If $1 - c = B(0)$ then $q_B = 0$. Since $u_B < 0$, we use (C7) of Lemma C to obtain statement [4].

That concludes the proof. □

3. Proof of Proposition 2

Recall that q_B and q_{ab} were introduced in Lemma A to denote the newsvendor's smallest optimal *ex ante* order under the CDF B and the CDF T_{ab}, respectively. Analogously, let S_B and S_{ab} denote, respectively, the average *ex post* sales under the CDF B and the CDF T_{ab} when the *ex ante* quantities are q_B and q_{ab}, i.e.,

$$S_B = E_B \min(q_B, D), \ S_{ab} = E_{T_{ab}} \min(q_{ab}, D).$$

From Lemma A we know that

$$q_{ab} = a + bq_B.$$

So we have:

$$
\begin{aligned}
S_{ab} &= \int \min(a + bq_B, D) dB \left(\frac{D - a}{b} \right) \\
&= \int \min(a + bq_B, a + bz) \ dB(z) \\
&= \int [a + b\min(q_B, z)] \ dB(z) \\
&= a + bS_B
\end{aligned}
$$

It follows that

$$\hat{S}(\eta) = E_{y \in Y^\eta} \ [a(y, \eta) + b(y, \eta) \cdot S_B].$$

Since, as noted in the proof of Lemma B, $E_{y \in Y^\eta} \ a(y, \eta) = \bar{D}$, we obtain

$$\hat{S}(\eta) = \bar{D} + V(\eta) \cdot S_B.$$

Then (C1) of Lemma C tells us that we indeed have

$$U(\eta') > U(\eta) \iff \hat{S}(\eta') > \hat{S}(\eta)$$

30

if $S_B < 0$. But that is the case, since

$$S_B = \int_\infty^\infty \min(q_B, z)dB(z) < \int_\infty^\infty zdB(z) = 0.$$

That concludes the proof. □

4. Proving Proposition 3.

4.1 Extended versions of Lemma A and Lemma B.

We first need extended versions of Lemma A and Lemma B. They exploit the fact that Proposition 3's payoff function obeys condition $(1')$.

Lemma A$'$

Let $\mathbb{R}$ be the set of possible values of an action-chooser's quantity q. Let the action-chooser's payoff function $u(q, \theta)$ satisfy (6) in the statement of Proposition 3. Consider an IG who has a base CDF B on $\mathbb{R}$ with mean zero. The IG determines a location parameter a and a scale parameter $b > 0$, which define the posterior CDF T_{ab}. Let q_B denote the smallest maximizer of $E_B\, u(q, \theta)$ and let q_{ab} denote the smallest maximizer of $E_{T_{ab}}u(q, \theta)$. Let u_B denote $E_B\, u(q_B, \theta)$ and let $u_{T_{ab}}$ denote $E_{T_{ab}}\, u(q_{ab}, \theta)$. Then

(A1$'$)
$$q_{ab} = a + b \cdot q_B.$$

and

(A2$'$) for any function f satisfying (†) in $(1')$, we have $u_{T_{ab}} = f(b) \cdot u_B$.

Proof:

For a function f satisfying (†) in $(1')$, we have

(i)
$$E_B\, u(q, a + b\theta) = f(b) \cdot E_B u\left(\frac{q - a}{b}, \theta\right).$$

Thus

(ii)
$$E_{T_{ab}}\, u(q, \theta) = f(b) \cdot E_B\, u\left(\frac{q - a}{b}, \theta\right).$$

We repeat the argument in the first paragraph of the proof of Lemma A to obtain

(iii)
$$u_{T_{ab}} = \max_q E_{T_{ab}}\, u(q, \theta) = f(b) \cdot \max_q E_B\, u\left(\frac{q - a}{b}, \theta\right).$$

Since $f(b) > 0$, we then have (A2$'$). We now repeat the remaining part of the proof of Lemma A, replacing (vi) of that proof with

$$E_{T_{ab}} \, u(a + bq_B, \theta) = f(b) \cdot E_B \, u(q_B, \theta).$$

That establishes (A1$'$)   □.

To state Lemma B$'$, we need a new symbol. Given a function $f : \mathbb{R} \to \mathbb{R}$ and the IG index η, define

$$V_f(\eta) = E_{y \in Y^\eta} \, f(b(y, \eta)).$$

We again consider a payoff function $u(q, \theta)$ satisfying (6) in the statement of Proposition 3. The Lemma concerns V_f and the functions $\hat{q}, U$, where

$$\hat{q}(\eta) = E_{y \in Y^\eta} \, q_{a(y,\eta), b(y,\eta)}; U(\eta) = E_{y \in Y^\eta} \, u(q_{a(y,\eta), b(y,\eta)}, \theta),$$

and, as before, $q_{a(y,\eta), b(y,\eta)}$ is the smallest maximizer of $E_{T_{a(y,\eta), b(y,\eta)}} \, u(q, \theta)$. The symbols q_B, u_B have the same meaning as in Lemma A$'$.

Lemma B$'$

Let f be a function satisfying (†) in our payoff-function property (1$'$). For every pair (η', η), the following hold:

(B1$'$) $\hat{q}(\eta') - \hat{q}(\eta) = q_B \cdot (V_f(\eta') - V_f(\eta))$

(B2$'$) $U(\eta') - U(\eta) = u_B \cdot (V(\eta') - V(\eta)).$

Proof:

The proof again obtains the statements (i)-(iv) in the proof of Lemma B (with θ replacing D). Again, (iv) implies (B1$'$). Moreover, (i),(ii), and (A2$'$) imply (B2$'$).   □

The following additional Lemma will be used in proving Proposition 3.

Lemma D

Suppose that the support of the base CDF B is the interval $[v, w]$, where $v < w$, v may be $-\infty$, and w may be ∞. Suppose that B has a positive density at every point of its support. Consider a payoff function u which satisfies (6) in the statement of Proposition 3. Then there exists a function q^* on $\mathbb{R}^+$ such that for every $K > 0$, $q^*(K)$ is the unique maximizer of $E_B \, u(q, \theta)$ on $\mathbb{R}$. For all $K > 0$, the expected payoff $u_B = E_B \, u(q^*(K), \theta)$ is negative. The function q^* has the following properties.

(i) q^* is strictly increasing.

(ii) $\lim_{K\to\infty} q^* > 0, \lim_{K\to 0} q^* < 0$.

(iii) There exists a critical value K^* for which

$$q^*(K) \begin{cases} > 0 & \text{if and only if } K > K^* \\ = 0 & \text{if and only if } K = K^* \\ < 0 & \text{if and only if } K < K^*. \end{cases}$$

(iv) For $r = 2$, the critical value K^* equals one.

(v) $K^* = 1$ if B is symmetric around zero, i.e., for any θ we have $B(-\theta) = 1 - B(\theta)$.

Proof:

We show that there is indeed a unique maximizer $q^*(K)$ and that the function q^* satisfies **(i)** -**(v)**.

Step 1.

Define
$$\phi(q) \equiv E_B \, u(q, \theta).$$

The support of B is $[v, w]$, where $v < w$ and v may be $-\infty$ while w may be ∞. If v is finite, then any $q < v$ cannot maximize ϕ (since $q = v$ yields a higher value of ϕ), and if w is finite, then any $q > w$ cannot maximize ϕ (since $q = w$ yields a higher value of ϕ). So we can ignore all $q \notin [v, w]$. With a slight abuse of notation, we write

$$\begin{aligned} \phi(q) &= \int_v^q -|q - \theta|^r dB(\theta) + \int_q^w -K|q - \theta|^r dB(\theta) \\ &= -\left[\int_v^q (q - \theta)^r dB(\theta) + \int_q^w K \cdot (\theta - q)^r dB(\theta) \right]. \end{aligned}$$

Fixing K and differentiating with respect to q, we obtain

$$(\dagger) \qquad \phi'(q) = -r \cdot \left[\int_v^q (q - \theta)^{r-1} dB(\theta) - K \cdot \int_q^w (\theta - q)^{r-1} dB(\theta) \right].$$

This expression is negative for some q and positive for others.

To see this, first let v, w be finite. By assumption $v < q < w$. For $q = v$, the first integral in $(\dagger)$ equals zero and the second integral is positive, so $\phi'(v) > 0$. For $q = w$, the first integral is negative and the second integral equals zero, so $\phi'(w) < 0$.

Next, let $v = -\infty$. Since $\lim_{q\to-\infty} \int_{-\infty}^q (q - \theta)^{r-1} dB(\theta) = 0$, we have

$$\lim_{q\to-\infty} \phi'(q) > 0.$$

33

So there exists a $q > 0$ at which $\phi'(q) > 0$. As before, we also have $\phi'(w) < 0$.

A symmetric argument covers the case where v is finite but $w = \infty$. A combination of both arguments covers the case where $v = -\infty, w = \infty$.

We conclude that the first-order condition $\phi'(q) = 0$ has a solution. To argue its uniqueness, we now examine the second derivative of ϕ. We have

$$\phi''(q) = -r \cdot (r-1) \left[\int_v^q (q-\theta)^{r-2} dB(\theta) + \int_q^w K \cdot (\theta - q)^{r-2} dB(\theta) \right].$$

Each of the integrals is nonnegative and at least one of them is positive (since $v < q$ or $q < w$). Moreover $r > 1$. Hence

$$\phi''(q) < 0 \text{ at all } q.$$

So the solution to the first-order equation is unique.

Step 2.

In calculating ϕ' and ϕ'' in the previous step, we treated K as fixed. We now permit K to take any value in $I\!\!R^+$ and we let $G(q, K)$ denote $\dfrac{d\phi}{dq}$ at the pair (q, K). We have seen that for every K, there is a unique maximizer of ϕ. We now let $q^*(K)$ denote the maximizer.

Letting G_K, G_q denote first derivatives, and differentiating both sides of the equation $G(q^*(K), K) = 0$, we obtain

(††) $$G_q(q^*(K), K) \cdot q^{*'}(K) + G_K(q^*(K), K) = 0.$$

In Step 1 we showed that for any K, we have $\phi''(q) < 0$ at all q, or, in our new notation, $G_q < 0$. Hence (††) can be written (in abbreviated form)

$$q^{*'} = -\frac{G_K}{G_q}.$$

So we have shown that $q^{*'} > 0$ once we have shown that $G_K > 0$ (or $\phi'' > 0$ in the previous notation) when $q = q^*(K)$. Since K does not appear in the first integral in the expression for G (or ϕ') in (†), we obtain, when we differentiate with respect to K,

$$G_K = r \cdot \int_q^w (\theta - q)^{r-1} dB(\theta).$$

For $q = q^*(K)$, that integral is positive, since (as we observed above) the maximizer $q^*(K)$ cannot equal w. Thus G_k is indeed positive and **(i)** in our Claim is established.

Step 3.

We now establish **(ii)**.

Consider a fixed $q_0 < w$. We have

$$G(q_0, K) = -r \int_0^{q_0} (q_0 - \theta)^{r-1} dB(\theta)) + Kr \cdot \int_{q_0}^w (\theta - q_0)^{r-1} dB(\theta).$$

Note that the second term is linear and increasing in K, while the first term is negative and does not contain K. So there must exist a sufficiently large K, say K_0, at which $G(q_0, K_0) = 0$. That is equivalent to $q^*(K_0) = q_0$. Since q^* is strictly increasing, we have $q^*(K) > q^*(K_0) = q_0$ for every $K > K_0$. Thus we have

$$\lim_{K \to \infty} q^*(K) > q_0 \text{ for any } q_0 < w.$$

That implies in turn that

$$\lim_{K \to \infty} q^*(K) = w.$$

A similar proof shows that $\lim_{K \to 0} q^*(K) = v$.

Step 4.

Statement **(ii)**,which we have just established, implies **(iii)**.

Step 5.

The statement **(iv)** concerns the critical K when $r = 2$. To show **(iv)** , we first note that

$$
\begin{aligned}
G(q, 1) &= \frac{d}{dq} E_B \left([(-1) \cdot (q^2 - 2\theta q + \theta^2)] \right) \\
&= \frac{d}{dq} \left[(-q^2 + (2q) \cdot (E_B \theta) - E_B \theta^2 \right] \\
&= -2q + 2E_B \theta \\
&= -2q
\end{aligned}
$$

(since $E_B \theta = 0$ by assumption). The derivative vanishes at zero, so the maximizer of $G(q, 1)$ is $q^*(1) = 0$. Since q^* is strictly increasing, the critical value of K must be $K^* = 1$, as **(iv)** asserts.

Step 6.

Finally, we show **(v)**.

By assumption, B has mean zero and $B(-\theta) = 1 - B(\theta)$ for all $\theta \in [v, w]$. That implies that $w = -v$. So we have (using (†))

$$G(0, 1) = -r \left[\int_v^0 (-\theta)^{r-1} dB(\theta) - 1 \int_0^w \theta^{r-1} dB(\theta) \right].$$

Now define the variable $\tau = -\theta$. Since $w = -v$ and $dB(-\tau) = -dB(\tau)$, we have

$$\int_v^0 (-\theta)^{r-1} dB(\theta) = \int_{-v}^0 \tau^{r-1} dB(-\tau) = \int_0^w \tau^{r-1} dB(\tau) = \int_0^w \theta^{r-1} dB(\theta).$$

35

Thus $G(0,1) = 0$ and hence $q^*(1) = 0$. Since q^* is strictly increasing, the critical value of K must be $K^* = 1$, as **(v)** asserts. □

Proof of Proposition 3:

We present the proof for the case where the support of B has the property considered in Lemma D: it is an interval $[v, w]$, where $v < w$, v may be $-\infty$, and w may be ∞.

Lemma D tells us that there is a unique maximizer of $E_B\ u(q, \theta)$, namely $q^*(K)$. By assumption, B is not a one-point distribution. Hence the probability that $\theta \neq q^*(K)$ is positive. But u, as defined in (+), is negative for all $\theta \neq q^*(K)$. It follows that $u_B < 0$, where u_B is defined as in Lemma A'.

Since $u_B < 0$, we can now repeat the argument that established statement **[1]** in Proposition 1. That gives us statement **[1]** of the present Proposition.

Using **(iii)** in Lemma D, and the argument that established statement **[2]** in Proposition 1, we obtain statement **[2]** of the present Proposition.

Using **(iii)** in Lemma D, and the arguments that established statements **[3]** and **[4]** in Proposition 1, we obtain statement **[3]** of the present Proposition.

Statements **[4]** and **[5]** of the present proposition are the same as (iv) and (v) of Lemma D. □

5. Proof of Proposition 4.

Just before the statement of this proposition the symbols $q_B, q_{ab}, q_{a(y,\eta),b(y,\eta)}$ were given new meanings. The proof of Proposition 2 uses those symbols. We now repeat the proof of Proposition 2, giving the symbols their new meanings. That establishes Proposition 4. □

6. Proof of the claim at the end of Section 5.1, concerning Condition ($*$) in Proposition 3.

The claim concerns the case where the IG, in each of his structures, partitions $\mathcal{D} = [0,1]$ into m subintervals whose widths need not be equal. If $\eta' > \eta$, and if $w' = (w'_1, \ldots, w'_m)$ (the subinterval-width vector for the structure $I_{\eta'}$) majorizes $w = (w_1, \ldots, w_m)$ (the subinterval-width vector for the structure I_η), then — we claim — $V_r(\eta') > V_r(\eta)$ if and only if $V(\eta') > V(\eta)$. (Here $V(\eta)$ is the average of the widths and $V_r(\eta)$ is the average of the rth power of the widths).

To prove the claim we use *Schur convexity*. (See Marshall and Olkin (1979)). A function ℓ on the width vectors is Schur-convex if $\ell(w^*) \geq \ell(w^{**})$ whenever w^* majorizes w^{**}. Recall that the probability of a given interval equals its width. Define $g(w)$ to be average scale and define $h(w)$ to be the average of the rth power of scale, i.e., $g(w) = (w_1)^2 + \cdots + (w_m)^2$ and $h(w) = (w_1)^{r+1} + \cdots + (w_m)^{r+1}$. Both the function g and the function h are symmetric and convex, and it is known that a symmetric convex function is Schur-convex. Since w' majorizes w, we can conclude that when we go from η to η', the functions g and h indeed move in the same direction, and thus V and V_r move in the same direction. That establishes the claim.

REFERENCES

Agrawal, V., S. Seshadri. 2000. Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem. *Manufacturing and Service Operations Management* **2**(4) 410-423.

Anand, K.S. 2009. Can information and inventories be complements? Working paper, Wharton School, University of Pennsylvania, Philadelphia, PA.

Anand, K.S., M. Goyal. 2009. Strategic information management under leakage in a supply chain. *Management Science* **55**(3) 438-452.

Bell, D. E. 1983. Risk premiums for decision regret. *Management Science* **29**(10) 1156-1166.

Bensoussan, A., M. Çakanyildirim, Q. Feng, S.P. Sethi. 2009. Optimal ordering policies for stochastic inventory problems with observed information delays. *Production and Operations Management.* **18**(5) 546-549.

Bergen, M., Iyer, A.V. 1997. Quick response in manufacturer-retailer channels. *Management Science* **43**(4) 559-570.

Blackwell, D. 1951. Comparison of experiments. J. Neyman, ed. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability.* University of California Press, Berkeley, 93-102.

Blackwell, D. 1953. Equivalent comparison of experiments. *Annals of Mathematical Statistics* **24** 265-273.

Boyaci, T., Özer, Ö. 2010. Information acquisition for capacity planning. *Operations Research* **58**(5) 1328-1349.

Buzacott, J. A., J.G. Shantikumar. 1994. Safety stock versus safety time in MRP controlled production systems. *Management Science* **40**(10) 1678-1689.

Cachon, G.P., M. Fisher. 1997. Campbell Soups' continuous replenishment program: evaluation and enhanced inventory decision rules. *Production and Operations Management* **6**(3) 266-276.

Cachon, G.P., M. Fisher. 2000. Supply chain inventory management and the value of shared information. *Management Science* **46**(8) 1032-1048.

Choi, T-M, D. Li, H. Yan. 2003. Optimal two-stage ordering policy with Bayesian information updating. *Journal of the Operational Research Society* **54** (8) 846-859.

Choi, S., A. Ruszczynski. 2008. A risk-averse newsvendor with law invariant coherent measures of risk. *Operations Research Letters* **36** 77-82.

Chu, L.Y., J.G. Shanthikumar, Z.J.M. Shen. 2008. Solving operational statistics via a Bayesian analysis. *Operations Research Letters* **36** 110-116.

Dudley, Leonard, Pierre Lasserre. 1989. Information as a substitute for inventories. *European Economic Review* **33**(1) 67-88.

Eeckhoudt, L., C. Gollier, R. Schlesinger. 1995. The risk-averse (and prudent) newsboy. *Management Science* **41**(5) 786-794.

Gallego, G., Ö. Özer. 2001. Integrating replenishment decisions with advance demand information". *Management Science* **47**(5) 1344-1360.

Gayon, J.P, S. Benjaafar, F. de Vericourt. 2009. Using imperfect demand information in production-inventory systems with multiple customer classes. *Manufacturing and Service Operations Management* **11**(1) 128-143.

Gaur, V., A. Giloni, S. Seshadri. 2005. Information sharing in a supply chain under ARMA demand. *Management Science* **51**(6) 961-969.

Gurnani, H., C. Tang. 1999. Note: optimal ordering decisions and demand forecast updating. *Management Science* **45**(10) 1456-1462.

Gerchak, Y., D. Mossman. 1992. On the effect of demand randomness on inventories and costs. *Operations Research* **40**(4) 804-807.

R. Hariharan, P. Zipkin. 1995. Customer-order information, leadtimes and inventories. *Management Science* **41** 1599-1607.

Hayashi, T. 2008. Regret aversion and opportunity dependence. *Journal of Economic Theory* **139**(1) 242-268.

Ho, T-H, N. Lim, T. Cui. 2010. Reference dependence in multilocation newsvendor models: a structural analysis. *Management Science* **56**(11) 1891-1910.

Iyer, G., C. Narasimhan, R. Niraj. 2007. Information and inventory in distribution channels. *Management Science* **53**(10) 1551-1561.

Jain, A., K. Moinzadeh. 2005. A supply chain model with reverse information exchange. *Manufacturing and Service Operations Management* **7**(4) 360-378.

Karaesmen, F., J. A. Buzacott, Y. Dallery. 2002. Integrating advance order information in production ontrol, *IIE Transactions* **34** 649-662.

F. Karaesmen, G. Liberopoulos, Y. Dallery. 2004. The value of advance demand information in production/inventory systems. *Annals of Operations Research* **126** 135-157.

H. S. Lau. 1980. The newsboy problem under alternative optimization objectives. *The Journal of the Operational Research Society* **31**(6) 525-535.

Lau, A.H., H.S. Lau. 2002. The effects of reducing demand uncertainty in a manufacturer-retailer channel for single-period products. *Computers and Operations Research* **29** 1583-1602.

Lee, H.L., K.T. So, C.S. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Science* **46**(5) 626-643.

Li, L., H. Zhang. 2008. Confidentiality and information sharing in supply chain coordination. *Management Science* **54**(8) 1467-1481.

Li, Q. and D. Atkins. 2005. On the effect of demand randomness on a price/quantity setting firm. *IEE Transactions* **37** 1143-1153.

Lu, Y., J. Song, D. Yao. 2003. Order fill rate, leadtime variability, and advance demand information in an assemble-to-order system. *Operations Research* **51**(2) 292-308.

Marschak, J., Miyasawa, K. 1968. Economic comparability of information systems. *International Economic Review* **9** 137-174.

Marschak, J., Radner R. 1972. *Economic Theory of Teams.* Yale University Press, New Haven.

Marschak, T., J.G. Shanthikumar, J. Zhou. 2013. Are information gathering and producing complements or substitutes? Working paper, SSRN eLibrary 2368076.

Marshall, A. and I. Olkin. 1979. *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York.

Milgrom, Paul, J. Roberts. 1988. Communication and inventory as substitutes in production. *Scandinavian Journal of Economics* **90**(3) 275-289.

Özer, Ö. 2003. Replenishment strategies for distribution systems under advance demand information. *Management Science* **49**(3) 255-272.

Özer, Ö, W. Wei. 2004. Inventory control with limited capacity and advance demand information. *Operations Research* **52**(6) 988-1000.

Perakis, G., G. Roels. 2008. Regret in the newsvendor model with partial information. *Operations Research* **56**(1) 188-203.

Prasad, A., K. Stecke, X. Zhao. 2011. Advance selling by a newsvendor retailer. *Production and Operations Management* **20**(1) 129-142.

Raghunathan, S. 2001. Information sharing in a supply chain: a note on its value when demand is nonstationary. *Management Science* **47**(4) 605-610.

Ridder, A., E. Van der Laan, M. Salomon. 1998. How larger demand variability may lead to lower costs in the newsvendor problem. *Operations Research* **46**(8) 934-936.

Schweitzer, S., G. Cachon. 1993. Decision bias in the newsvendor problem with a known demand distribution: experimental evidence. *Management Science* **46**(3) 404-420.

Song, J., F. Zipkin. 1993. Inventory control in a fluctuating demand environment. *Operations Research* **41**(1) 351-370.

Song, J.. F. Zipkin. 2012. Newsvendor problems with sequentially revealed demand information. *Naval Research Logistics Quarterly* **59** 601-612.

Song, J. 1994. The effect of leadtime uncertainty in a simple stochastic inventory model *Management Science* **40**(5) 603-613.

Song, J., H. Zhang, Y. Hou, M. Wang. 2010. The effect of lead time and demand uncertainties in (r,q) Inventory Systems. *Operations Research* **58**(1) 68-80.

Toktay, L., L. Wein. 2001. Analysis of a forecasting-production-inventory system with stationary demand. *Management Science* **47**(9) 1268-1281.

M. Xu, Y. Chen, X. Xu. 2010. The effect of demand uncertainty in a price-setting newsvendor model. *European Journal of Operational Research* **27**. 946-95.

H. Zhang. 2002. Vertical information exchange with duopoly retailers. **Production and Operations Management 11**(4), 531-546.