

On the Robustness of ‘Good News and Bad News’*

Christopher P. Chambers[†] Paul J. Healy[‡]

Abstract

Paul Milgrom (*The Bell Journal of Economics*, 12(2): 380–391) showed that if the strict monotone likelihood ratio property (MLRP) does *not* hold for a signal structure then there exists *some* non-degenerate prior and a pair of signals where the higher-signal posterior does not stochastically dominate the lower-signal posterior. We show here that for *any* non-degenerate prior with bounded support there exists an additive signal structure (satisfying other natural properties) and pair of signals such that a higher signal implies a *lower* posterior. Thus, higher signals can be “bad” news when the prior is fixed but the signal structure is not.

Keywords: Signal extraction, Bayes’s rule, MLRP, stochastic dominance, updating
JEL: C11, C60, D81, D84

*The authors thank Jim Peck for his valuable comments.

[†]Division of the Humanities and Social Sciences, California Institute of Technology, 1200 E. California Boulevard, MC 228-77, Pasadena, CA, 91125. Phone: (626) 395-3559 Email: chambers@hss.caltech.edu.

[‡]Corresponding author. Department of Economics, The Ohio State University, 1945 North High Street, Columbus, OH 43210. Phone: (614) 247-8876. Fax: (614) 292-3906. Email: healy.52@osu.edu.

1 Introduction

Suppose that a single-dimensional parameter is observed with noise, where the noisy observation is considered to be an informative *signal* of the parameter. For two possible signals z and z' with $z < z'$ when is it possible to say that the posterior distribution of the parameter after observing z' first-order stochastically dominates the distribution after observing z , regardless of the structure of the prior distribution of the parameter?¹ Milgrom's celebrated "good news, bad news" result (Milgrom, 1981) establishes that it is necessary and sufficient that the family of conditional distributions (technically, the family of conditional distributions of the parameter obtained after conditioning on an observation of the signal—hereafter referred to as a *signal structure*) satisfy the strict *monotone likelihood ratio property* (MLRP).²

Although Milgrom's characterization is extremely useful, its contrapositive is quite weak; knowing that a signal structure *fails* the strict MLRP only implies that there exists *some* prior distribution and pair of signals such that the posterior generated by the higher signal does not stochastically dominate the posterior generated by the lower signal. In fact, since stochastic dominance does not completely order distributions, it may be that the higher-signal posterior is simply incomparable to the lower-signal posterior. It need not be true that the lower-signal posterior stochastically dominates the higher-signal posterior.

Upon inspecting Milgrom's proof it is easy to see how the contrapositive can be strengthened. Using Milgrom's arguments, one can show that if a signal structure fails the strict MLRP then there is *some* prior distribution and pair of signals such that the lower-signal posterior stochastically dominates the higher-signal posterior. This is done by constructing a prior that places mass on only two points. For priors with larger support, however, Milgrom's arguments are ambiguous about whether or not signal monotonicity will be reversed when the signal structure fails the strict MLRP.

In this note we obtain a stronger conclusion about the failure of the strict MLRP by taking the prior distribution as fixed and allowing for (limited) flexibility in the signal structure. Formally, we show that for *any* non-degenerate prior distribution with bounded support there exists some signal structure (obviously failing the strict MLRP) and some pair of signals such that the lower-signal posterior stochastically dominates the higher-signal posterior. In fact, the signal structure we identify to generate this result is quite well-behaved. For any

¹A cumulative distribution function F strictly first-order stochastically dominates G if $F(x) \leq G(x)$ for all x with strict inequality for some x .

²If x and x' are two realizations of the parameter, z and z' are two realizations of the noisy signal, and $f(z|x)$ is the density of the conditional distribution of the signal, then the strict MLRP requires that $f(z'|x')f(z|x) > f(z'|x)f(z|x')$ for all $x' > x$ and $z' > z$.

parameter realization, its density is continuous, bounded, symmetric about the realization of the underlying variable, and quasiconcave. Furthermore, it is an *additive* signal structure: the signal is the parameter value plus some noise term whose distribution is independent of the distribution of the parameter value.

To interpret this result, consider a modeler studying some variable that is observed with noise. The modeler fixes a particular prior distribution for his beliefs about the variable but, in the interest of robustness, only assumes that the signal structure satisfies certain properties (rather than assuming a specific functional form). Specifically, suppose he assumes that any additive signal structure with a mean-zero, symmetric, and quasiconcave density is admissible.³

A direct application of Milgrom's result does not indicate to us whether monotonicity (in the sense of first order stochastic dominance) with respect to the signal observation is guaranteed for our fictitious modeler; for each admissible signal structure failing strict MLRP, it may or may not be that the chosen prior is one for which monotonicity will fail. Even if monotonicity does fail, it may not be the case that a higher signal leads to a lower posterior; the posteriors may be incomparable.

Applying our result, however, leads to a more pessimistic conclusion: as long as the prior chosen has bounded support then there will always be an admissible signal structure and pair of signals such that monotonicity is reversed. The lower signal will generate a strictly higher posterior. Our result therefore suggests that when worrying about signal monotonicity one cannot stray 'too far' from the MLRP even when the prior is fixed.

We argue that fixing a prior and allowing for flexibility in the signal structure is a natural approach in many applications. The noise inherent in a signal often represents components of the underlying process that are not well understood or not properly captured by existing research. Strong assumptions about the structure of the noise (such as MLRP) may lead to models that are misspecified. The prior, on the other hand, is usually thought of as subjective (as in Bayesian statistics or decision theory) and therefore taken to be a fixed distribution.⁴

Note that requiring an additive signal structure makes our theorem *stronger* rather than *weaker*. We demonstrate the *existence* of a particular signal structure, so restricting this

³In a related paper Chambers and Healy (2007) we show that with this family of noise distributions the posterior mean lies between the prior mean and the signal observation if and only if the prior distribution is symmetric and quasiconcave.

⁴A subset of the literature on robust Bayesian analysis takes exactly this approach; see Berger (1994) for a survey.

signal structure to lie in a particular set makes the theorem more difficult to prove as we have a much smaller set of signal distributions available to demonstrate the result. This makes the pessimistic conclusion for our fictitious modeler all the more troublesome. Clearly, any larger set of families of signal distributions that he considers admissible will result in the same unfortunate conclusion.

The intuition of our proof is simple. For any prior distribution with bounded support consider a mean-zero error distribution whose support is significantly larger than the support of the prior (though still bounded). If this error distribution has sufficiently “fat” tails then any extremely large positive observation z' is likely due to a very large error term, indicating a relatively small parameter value. For a less extreme observation z that falls in the support of the prior it becomes more likely that the observation is indicative of a large parameter value. By carefully constructing the noise distribution one can guarantee that the posterior after observing z' is stochastically dominated by the posterior after observing z . In fact, the construction of the noise distribution needs only to depend on the support of the prior. For the case where the prior has support on $[-10, 10]$ the constructed noise distribution is shown in Figure 1. With this error term the posterior after observing $z' = 30$ is stochastically dominated by the posterior after observing $z = 10$ *for any* prior with support on $[-10, 10]$.

Our proof relies heavily on the support of the prior being bounded. This assumption cannot be dispensed of through standard limiting arguments, and we have yet to find a conditional distribution for unbounded priors that generates the same sort of monotonicity reversal as we have with bounded priors. Thus, the unbounded case remains an open question.

2 The model and main result

Given are two real-valued random variables, X and Z . In the interest of simplicity, we refrain from defining the underlying probability space on which these random variables are defined. X represents the distribution of some economically relevant parameter and Z is interpreted as a signal of this parameter. A realization of X is denoted x ; likewise, a realization of Z is denoted z . The family of distributions obtained by conditioning Z on $X = x$ for each x is referred to as a signal structure.⁵

X is said to be bounded if there exist $a, b \in \mathbb{R}$ for which the probability that X lies in $[a, b]$

⁵When x is outside the support of the prior, let the distribution of Z conditional on x be any arbitrary distribution.

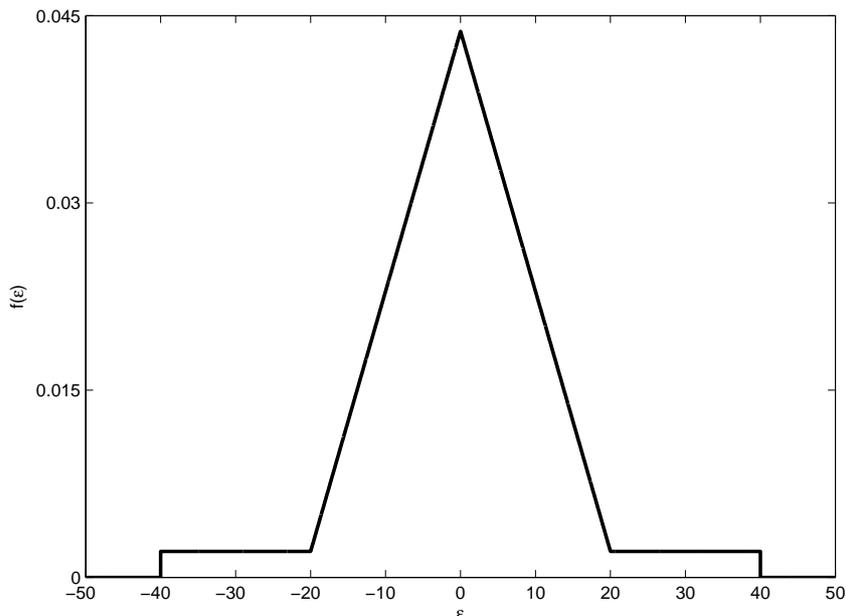


Figure 1: An example of the noise distribution used in the proof.

is equal to one. Cumulative distribution functions of random variables are denoted F with the appropriate subscript, whereas densities are denoted f . Thus, $F_{Z|x}$ is the cumulative distribution function of Z conditional on $X = x$, for example.

In our analysis it is convenient to define the random variable $\tilde{\varepsilon} := Z - X$. The signal structure obtained by Z is said to be additive if $\tilde{\varepsilon}$ is independent of X . In particular, this means that $Z = X + \tilde{\varepsilon}$ for $\tilde{\varepsilon}$ independent of X .

For completeness, we reproduce the contrapositive of Milgrom's original theorem here.

Theorem (Milgrom 1981). *If a family of conditional density functions $f_{Z|x}$ does not have the strict MLRP then there exists some non-degenerate prior distribution F_X and two signals $z' > z$ such that the posterior $F_{X|z'}$ does not first-order stochastically dominate $F_{X|z}$.*

As mentioned above, inspection of Milgrom's proof leads to a slightly stronger version of this result.

Corollary (Milgrom 1981). *If a family of conditional density functions $f_{Z|x}$ does not have the strict MLRP then there exists some non-degenerate prior distribution F_X (which puts mass on only two points) and two signals $z' > z$ such that the posterior $F_{X|z'}$ is strictly first-order stochastically dominated by $F_{X|z}$.*

The following theorem shows how signal monotonicity can be reversed for *any* non-degenerate, bounded prior if the modeler cannot commit to a particular noise (or conditional) distribution.

Theorem. *For every non-degenerate, bounded X there exists a signal structure Z and two signal realizations $z' > z$ such that $F_{X|z'}$ is strictly first-order stochastically dominated by $F_{X|z}$. Furthermore, Z can be chosen to have the following properties: i) Z is an additive signal structure ii) $\tilde{\varepsilon} := Z - X$ is mean-zero, symmetric, quasiconcave, and has bounded support.*

Proof. Let $[a, b] \subset \mathbb{R}$ (with $a < b$) be the support of X and let $Z = X + \tilde{\varepsilon}$ where $\tilde{\varepsilon}$ is a noise term distributed according to the density

$$f_{\tilde{\varepsilon}}(\varepsilon) = \begin{cases} \frac{1}{4(b-a)+(b-a)^2} & \text{for } \varepsilon \in [-2(b-a), -(b-a)] \cup [(b-a), 2(b-a)] \\ \frac{1}{4(b-a)+(b-a)^2}(b-a+1+\varepsilon) & \text{for } \varepsilon \in (-(b-a), 0] \\ \frac{1}{4(b-a)+(b-a)^2}(b-a+1-\varepsilon) & \text{for } \varepsilon \in (0, (b-a)). \end{cases}$$

(An example of this distribution is shown in Figure 1.) Let $Z = X + \tilde{\varepsilon}$ and consider $z = b$ and $z' = 2b - a$, which are two feasible realizations of Z such that $z < z'$. Fix any $w \in [a, b]$ and note that the cumulative distribution function of $X|z$ is given by

$$F_{X|z}(w) = \frac{\int_a^w (x - a + 1) dF_X(x)}{\int_a^b (x - a + 1) dF_X(x)}. \quad (1)$$

Moreover, note that $X|z'$ is simply distributed the same as X , so that $F_{X|z'} \equiv F_X$.

Separately integrating the numerator and denominator of (1) by parts and rearranging, we obtain

$$\begin{aligned} F_{X|z}(w) &= \frac{(w - a + 1)F_X(w) - \int_a^w F_X(x) dx}{(b - a + 1) - \int_a^b F_X(x) dx} \\ &= \frac{F_X(w) - \int_a^w [-F_X(w) + F_X(x)] dx}{1 - \int_a^b [-1 + F_X(x)] dx} \\ &= \frac{F_X(w) + \int_a^w [F_X(w) - F_X(x)] dx}{1 + \int_a^b [1 - F_X(x)] dx}. \end{aligned} \quad (2)$$

Clearly, if $F_X(w) = 0$ then this expression calculates to 0 at w and hence $F_{X|z}(w) \leq F_{X|z'}(w)$, consistent with $F_{X|z}$ stochastically dominating $F_{X|z'}$. If $F_X(w) = 1$ then obviously $F_{X|z}(w) \leq F_{X|z'}(w)$ since $F_{X|z'}(w) = 1$.

Finally, consider the case where $F_X(w) \in (0, 1)$. For these values of w the following is true of the numerator of (2):

$$\begin{aligned} F_X(w) + \int_a^w [F_X(w) - F_X(x)]dx &= F_X(w) \left(1 + \int_a^w \left[1 - \frac{F_X(x)}{F_X(w)} \right] dx \right) \\ &\leq F_X(w) \left(1 + \int_a^w [1 - F_X(x)]dx \right) \\ &\leq F_X(w) \left(1 + \int_a^b [1 - F_X(x)]dx \right) \end{aligned}$$

If $w > a$ then the first inequality is strict since $F_X(w) < 1$. If $w = a$ then the second inequality is strict because $b > a$.⁶ Dividing by the term in parentheses, we thus establish that

$$\frac{F_X(w) + \int_a^w [F_X(w) - F_X(x)]dx}{1 + \int_a^b [1 - F_X(x)]dx} < F_X(w).$$

Recalling equation (2) and the fact that $F_{X|z'} \equiv F_X$, the above inequality implies $F_{X|z}(w) < F_{X|z'}(w)$. Therefore, $F_{X|z}$ strictly first order stochastically dominates $F_{X|z'}$. \square

The following key points are important.

- Requiring our signal structure to be additive and to satisfy additional conditions results in a much stronger theorem than if no such conditions were required. If the signal structure were not required to satisfy any conditions, setting $Z = -X$ would trivially establish our result.
- Our result is not implied by, nor does it imply, Milgrom's result. Nor do simple modifications of either result imply the other. To be clear, the difference lies in the quantification. Milgrom's result shows that *for any* signal structure, *there exists* a prior distribution generating a certain property (failure of FOSD). Our result has the quantifiers reversed, so that *for any* prior distribution, *there exists* a signal structure generating a certain property (the reversal of FOSD). This distinction in quantification is critical.
- The corollary of Milgrom's result given above generates a reversal of signal monotonicity using a particular prior distribution with a *two-point* support. Focusing on

⁶Recall that this case assumes $F_X(w) > 0$, so $w = a$ implies a point mass at a . Since $b > a$ it cannot be that $F_X(w) = 1$.

priors with two-point supports necessarily strengthens the contrapositive of Milgrom’s original theorem because the FOSD relation restricted to the family of distributions which have the same two-point supports is complete. Thus, we emphasize the point alluded to in the previous bullet: Our theorem holds for *any* prior distribution which is non-degenerate and has bounded support—not just those whose support has only two points.

- Although our signal structure obviously must fail the strict MLRP (see below for verification of this fact), we argue that it is “natural” in most other respects. In particular, symmetry around x and quasiconcavity of the density imply that signals are unbiased and signals closer to x are more likely than signals farther from x .

We now verify that for any X with bounded support, the conditional distribution used in the proof violates the strict MLRP. Let the support of X be $[a, b]$ with $a < b$ and consider $x = a$, $x' = b$, $z = b$, and $z' = 2b - a$. The strict MLRP requires that

$$f_{Z|x}(z)f_{Z|x'}(z') > f_{Z|x}(z')f_{Z|x'}(z),$$

or, using the fact that $\tilde{\varepsilon} = Z - X$ and substituting in the above values of x , x' , z , and z' ,

$$f_{\tilde{\varepsilon}}(b-a)f_{\tilde{\varepsilon}}(b-a) > f_{\tilde{\varepsilon}}(2(b-a))f_{\tilde{\varepsilon}}(0).$$

This expression evaluates to

$$\left(\frac{1}{4(b-a) + (b-a)^2}\right)^2 > \left(\frac{1}{4(b-a) + (b-a)^2}\right) \left(\frac{1 + (b-a)}{4(b-a) + (b-a)^2}\right),$$

but the right-hand side is strictly larger (since $b > a$), so the strict MLRP is violated.

References

- BERGER, J. O. (1994): “An Overview of Robust Bayesian Analysis,” *Test*, 3, 5–124.
- CHAMBERS, C. P., AND P. J. HEALY (2007): “On the Robustness of Good News and Bad News,” Ohio State University working paper.
- MILGROM, P. R. (1981): “Goods News and Bad News: Representation Theorems and Applications,” *The Bell Journal of Economics*, 12, 380–391.