

Reversals of Signal-Posterior Monotonicity for Any Bounded Prior*

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 conditional distribution then there exists *some* non-degenerate prior and pair of signals where the higher-signal posterior does not stochastically dominate the lower-signal posterior. We show that for *any* non-degenerate prior with bounded support there exists a conditional distribution (satisfying several natural properties) and pair of signals such that the lower signal’s posterior stochastically dominates that of the higher signal. Thus, for every bounded prior, higher signals may represent strictly “worse” news.

Keywords: Signal extraction, Bayes’s rule, MLRP, stochastic dominance, updating

JEL: C11, C60, D81, D84

*The authors thank Jim Peck, Simon Grant, and two anonymous referees for their valuable comments.

[†]Department of Economics, University of California, San Diego, MC0508 9500 Gilman Drive, La Jolla, CA 92093. Phone: (858) 534-2988. Email: cpchambers@ucsd.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.

The classic “good news, bad news” result of Milgrom (1981) shows that the strict monotone likelihood ratio property (MLRP) is both necessary and sufficient for higher signals of a noisy random variable to be “good news”, in the sense of first-order stochastic dominance. More formally, suppose Z is a noisy signal of X , where X is distributed according to F , and conditional on $X = x$, Z is distributed according to G_x , with density g_x . The family of all such conditional distributions is denoted $\{G_x\}$. This family satisfies the strict MLRP if, for all $z'' > z'$, the likelihood ratio $g_x(z'')/g_x(z')$ is increasing in x .¹ Denote the unconditional distribution of Z by G , and the distribution of X conditional on $Z = z$ by F_z . Using this notation, Milgrom’s result tells us that $F_{z''}$ first-order stochastically dominates $F_{z'}$ for all $z'' > z'$ *independently of F* if and only if the family $\{G_x\}$ satisfies the strict MLRP. The result is compelling, as we tend to think higher values of a noisy signal should be “good news” about the underlying parameter. Milgrom’s result tells us exactly when this is the case.

According to Milgrom’s result, a failure of the MLRP on $\{G_x\}$ does not preclude the possibility that, *for some F* , $F_{z''}$ first-order stochastically dominates $F_{z'}$ for all $z'' > z'$. In fact, it merely demonstrates that *there exists* some F , and a pair $z'' > z'$ for which $F_{z''}$ does not first-order stochastically dominate $F_{z'}$. This is not the same as saying that $F_{z'}$ first-order stochastically dominates $F_{z''}$, so this does not imply that a higher signal necessarily leads to “bad news”. It may depend on which prior is chosen.

Here, we ask if it is possible that a failure of the MLRP can lead to an “extreme” failure of Milgrom’s result, in the sense that a higher signal realization can lead to “bad news”, regardless of the prior. When the prior has a known, bounded support, we show that in fact it can, and we do so with a signal structure that seems reasonably “close” to satisfying the MLRP. Specifically, we choose $Z = X + \tilde{\varepsilon}$, where $\tilde{\varepsilon}$ is independent of X , unimodal, and symmetric.² Thus, higher values of x lead to higher signal distributions for Z , in the sense of stochastic dominance. Given some finite support $[a, b]$ for the prior, we show that there exists $\tilde{\varepsilon}$ and a pair $z'' > z'$ such that for any non-degenerate F whose support lies in $[a, b]$, $F_{z'}$ first-order stochastically dominates $F_{z''}$. Thus, the higher signal realization is “bad news”, no matter what the prior.

The intuition of our proof is simple. For any prior distribution with bounded support, consider a symmetric and unimodal signal distribution with mean equal to the parameter realization and whose support is significantly larger than the support of the prior (though still bounded). Thus, the signal equals the underlying parameter realization plus a high-variance, mean-zero ‘noise’ variable. If this error distribution has sufficiently ‘fat’ tails,

¹Definition 1 below gives a slightly more precise definition.

²Unimodality is equivalent to requiring that the density function be quasiconcave.

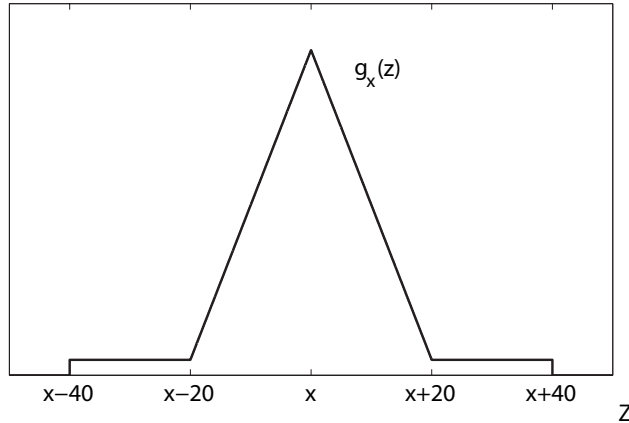


Figure 1: An example of the conditional density used in the proof.

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' stochastically dominates 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 conditional distribution (for any x) is shown in Figure . With this conditional distribution, the posterior after observing $z' = 10$ stochastically dominates the posterior after observing $z'' = 30$ for any prior with support on $[-10, 10]$.

Our proof relies heavily on the support of the prior being bounded. Our theorem does not hold if the prior is the (improper) uniform distribution over the entire real line. With this prior and a signal that is a mean-preserving spread, the signal realization simply shifts the location of the posterior distribution. Higher signals shift the entire posterior to the right, and so signal-posterior monotonicity is restored. Whether our result holds for integrable unbounded priors remains an open question.³

More carefully, given are a real-valued random variable X with cumulative distribution F and, for each realization x of X , a conditional random variable $Z|x$ with distribution G_x .⁴ Here, X represents some economically relevant parameter, and $Z|x$ a random signal of

³Our current proof uses conditional distributions with bounded support; if the prior were unbounded then a conditional distribution with bounded support would not generate the necessary reversal. Limiting arguments are problematic because the space of probability measures over the real line is not compact in the weak* topology, so even if a sequence of bounded priors converges to an unbounded prior, the required conditional distributions and signals need not converge.

⁴In the interest of simplicity, we refrain from defining the underlying probability space on which these

that parameter. Each G_x is assumed to have a well-defined density function g_x , and typical realizations are denoted by z .⁵ The family of conditional distributions is $\{G_x\}$, and the family of conditional densities is $\{g_x\}$.

A random variable X (and its distribution F) is said to be *bounded* if there exists some $a, b \in \mathbb{R}$ for which the probability that X lies in $[a, b]$ is equal to one. The *support* of X is the smallest such interval. X is *degenerate* if there is some $a \in \mathbb{R}$ such that $F(a) = 1$ and $F(b) = 0$ for all $b < a$; it is *non-degenerate* otherwise. If the conditional distributions are such that $Z|x - x$ is identical (in distribution) for every x , and if $E[Z|x] = x$ for each x , then we say that the signal forms an *independent additive signal* of X . This implies that the (unconditional) signal can be modeled as a random variable $Z = X + \tilde{\varepsilon}$ for some mean-zero random variable $\tilde{\varepsilon}$ that is independent of X .

The distribution F is referred to as the ‘prior’; upon observing any signal realization z a Bayesian observer’s posterior belief is given by the conditional distribution F_z , formed according to Bayes’s Law in the usual way.

For completeness, we state Milgrom’s sufficiency result here.

Definition 1 (MLRP). *A family of density functions $\{g_x\}$ has the strict monotone likelihood ratio property (MLRP) if $x'' > x'$ and $z'' > z'$ imply $g_{x''}(z'') g_{x'}(z') > g_{x''}(z') g_{x'}(z'')$. Thus, for any $z'' > z'$, $g_x(z'')/g_x(z')$ is strictly increasing in x .*

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

Inspection of Milgrom’s proof leads to a slightly stronger version of this result.

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

The following theorem is our main result. It 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. *Fix any $a < b$. There exists a family of conditional density functions $\{g_x\}$ and two signal realizations $z'' > z'$ such that for all X whose support is $[a, b]$, $F_{z'}$ strictly first-order stochastically dominates $F_{z''}$. Furthermore, $\{g_x\}$ forms an independent additive signal, and each g_x is unimodal and symmetric.*

random variables are defined.

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

Proof. Let $[a, b] \subset \mathbb{R}$ (with $a < b$) be the support of X , set $d = b - a$, and for each $x \in [a, b]$ let g_x be given by

$$g_x(z) = \begin{cases} \frac{1}{4d+d^2} & \text{for } z \in [x-2d, x-d] \cup [x+d, x+2d] \\ \frac{1}{4d+d^2}(1+d+(z-x)) & \text{for } z \in (x-d, x] \\ \frac{1}{4d+d^2}(1+d-(z-x)) & \text{for } z \in (x, x+d). \end{cases}$$

Note that g_x has a mean of x , is symmetric, and unimodal for each x ; an example of this distribution is shown in Figure . Now consider $z' = b$ and $z'' = b + d$, which are two feasible realizations of Z such that $z'' > z'$. Fix any $w \in [a, b]$ and note that the posterior distribution on X given z' is equal to

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

Moreover, note that the posterior of X conditional on z'' is distributed the same as the prior, so that $F_{z''} \equiv F$.

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

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

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

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

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

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

$$F_{z'}(w) = \frac{F(w) + \int_a^w [F(w) - F(x)]dx}{1 + \int_a^b [1 - F(x)]dx} < F(w).$$

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

The following key points are important.

- As stated, requiring the signal to be an independent additive signal of X and to satisfy additional properties results in a much stronger theorem than if no such conditions were required. If the signal distribution were not required to satisfy any conditions, setting $Z = -X$ would establish our result trivially.
- 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* conditional distribution failing MLRP *there exists* a prior distribution generating a reversal of signal monotonicity. Our result has the quantifiers reversed: *for any* prior distribution *there exists* a (well-behaved) conditional distribution generating a reversal of signal monotonicity. Actually, this conditional distribution can be chosen as a function of the support only. 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 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 the conditional distribution used in our proof 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

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

that signals are unbiased and signals closer to x are more likely than signals farther from x .

By Milgrom's result, it must be that, for any bounded X , the family of conditional distributions used in the proof violates the strict MLRP. We now verify this fact directly, for completeness. Let the support of X be $[a, b]$ with $a < b$, set $d = b - a$, and consider $x' = a$, $x'' = b$, $z' = b$, and $z'' = b + d$. The strict MLRP requires that

$$g_{x'}(z')g_{x''}(z'') > g_{x'}(z'')g_{x''}(z'),$$

or, substituting in the above values of x' , x'' , z' , and z'' ,

$$g_a(b)g_b(b + d) > g_a(b + d)g_b(b).$$

This expression evaluates to

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

but the right-hand side is strictly larger, so the strict MLRP is violated.

Finally, we illustrate our theorem with a simple application borrowed from Milgrom (1981). We imagine an economy with one risky and one riskless asset. The risky asset's returns have density f . We have a collection of identical agents, each of whom possesses the same differentiable utility u . Each consumer is endowed with one unit of the risky asset and one unit of the riskless asset. By normalizing the price of the riskless asset to one, the price of the risky asset (in equilibrium) is given by

$$p = \frac{E[Xu'(1 + X)]}{E[u'(1 + X)]}.$$

Note in particular that by defining the "density" function $h(x) = f(x)\frac{u'(1+x)}{E[u'(1+X)]}$, we get that $p = \bar{E}[X]$, where \bar{E} is the expectation for p , taken with respect to h . If, instead, agents observe information in the form of a noisy signal Z before trading, then the equilibrium price of the risky asset is given by

$$p(z) = \frac{E[Xu'(1 + X)|z]}{E[u'(1 + X)|z]}.$$

According to Milgrom, this is the same as $p(z) = \bar{E}[X|z]$.

Milgrom's theorem implies that if the noisy signal satisfies the strict MLRP, then $p(z)$ is monotonically increasing in z . The point of our theorem is to show that there can be

very well-behaved signal structures (specifically, where $Z = X + \tilde{\varepsilon}$ and $\tilde{\varepsilon}$ is independent of X , unimodal, and symmetric) under which monotonicity can be reversed for some signals: There are two signals $z'' > z'$ where $p(z'') < p(z')$.

References

MILGROM, P. R. (1981): "Goods News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 12, 380–391.