The Proportionate Likelihood Ratio Property

Andrea Wilson*
Georgetown University, Economics

June 2, 2019

Abstract

The *monotone likelihood ratio property* (MLRP) is widely used in economics. Equivalent to Karlin's (1968) total positivity of order 2 (log-supermodularity), many of its applications follow from Karlin's *variation diminishing property* (VDP).

I characterize the *proportionate likelihood ratio property (PLRP)*, which also asks that the likelihood ratio grow at a log-supermodular rate, and is equivalent to Karlin's total positivity of order 3: I derive an easily checked differential test for the PRLP in general, and an easier one for location family signals.

While the MLRP preserves monotonicity, the PLRP preserves quasiconcavity and quasiconvexity, and also ensures that expectations of more concave (or convex) functions grow more concave (or convex). So while the MLRP sees application in the vertical economic models, where higher is better, the PLRP is useful in horizontal (variety) economic models, such as matching with the right partner.

I illustrate the usefulness of this theory by many examples, including proofs of a few classic information economics results. I end with a new, quick, and gentle proof of the VDP, hoping to make this theory accessible to a larger audience.

*Email: andrea.q.wilson@gmail.com. I thank Lones Smith for many helpful comments, and Luciano Pomatto for a suggestion that motivated Theorem 1. I am also grateful to Mavi for furry input.

1 Introduction

The monotone likelihood ratio property (MLRP) — or a log-supermodular signal kernel — is an essential tool for information economics. It ensures that higher signal realizations lead to higher expected payoffs whenever the payoff function is increasing in the state. Given an MLRP kernel, expectation of functions that single cross through zero share this property (Karlin and Rubin, 1956). This observation inspired the later *variation diminishing property (VDP)*, summarized in the text Karlin (1968). The VDP says that under determinant conditions on the stochastic kernel, called *total positivity* (TP), an expectation of a function has weakly fewer sign changes than the function does, and in the same order. For instance, totally positive of order 2 (TP-2) is equivalent to the MLRP (or log-supermodularity). Next, totally positive of order 3 (TP-3) imposes an additional determinant condition on the kernel, and preserves two-crossing properties of the sign of the function, etc. This note has two purposes: I flesh out TP-3 functions — both how to practically check this property and how useful it is in economics — and I then offer a new and simple proof of the variation diminishing property.

I flesh out properties of signal families in §2. I seek to characterize the *proportionate likelihood ratio property* (PRLP) on kernels that strengthens the MLRP, also requiring that the likelihood ratio grow at a log-supermodular rate. The PLRP is equivalent to TP-3, and can easily be checked given a differentiable kernel (Theorem 1). I then specialize these conditions to additive location family signals (Corollary 1). Such signal families often induce discrete signals using truncations — e.g., "take the high action above a cutoff". The MLRP works well with truncation because the indicator function is log-supermodular, and so multiplication preserves it. While the PLRP lacks this useful preservation property, Theorem 2 rescues it when the kernel density vanishes near the truncation point, as happens in some canonical families.

I then explore how expectations reflect the properties of stochastic kernels in §3. Proposition 1 carefully lays out the VDP, nuancing weak and strict versions — where the latter assumes strict TP determinant signs. The MLRP ensures that increasing functions have increasing expectations given the signal, and posterior beliefs stochastically ranked in signals. So inspired, Milgrom called this the favorable news' ranking of signals. Corollaries 2 and 3 offer parallel presentations of these and analogous PLRP properties —- that integration preserves quasiconcavity (and quasiconvexity), and that slope of posterior cdf's in the signal are MLRP. The strict PRLP ensures that the expectation of a nonconstant quasiconcave function has no flats. I suggestively call this signal ranking *harmonious news*, since it preserves the curvature of both hump-shaped and U-shaped functions. Whereas the MLRP applies to "vertical" or "quality" models, with monotone objectives, the PLRP is designed for "horizontal" or "variety' models, with hump-shaped objectives, such as in matching and voting. One might then wonder if PLRP yields parallel preservation results for concavity and convexity. In Figure 2, I show by example that this is false: an inverted quadratic may well have a nonconcave expectation. Instead, the PLRP implies that more concave functions have more concave expectations, and more convex functions have more convex expectations (Theorem 3).

In §4, I turn to the behavior of optimal actions. By monotone comparative statics theory, a scalar optimizer is monotone in a scalar parameter when the marginal return is monotone in that parameter. This same comparative statics conclusion holds when maximizing an expectation given an MLRP kernel, since that preserves the single crossing property. Consider next an expectation with respect to a PLRP kernel. Corollary 4 says that in this case, the optimizer is quasiconcave (or quasiconvex) in a parameter when the marginal return of the integrand is quasiconcave (respectively, quasiconvex).

Finally, §5 offers a quick new "peasant's proof" of the VDP, avoiding Karlin's involved appeals to matrix decomposition. To see that taking expectations reduces sign variation, I consider the contrapositive: If an expectation has n sign changes, then so too does the function. This is obvious with just one sign change. The expectation of a positive function is positive, and so an expectation can only change signs if the function does. More generally, I treat the portion of the function between each pair of adjacent crossing points as a single "unknown", and show that asking an expectation to change sign n times when the function does not is like solving n equations in fewer than n unknowns.

My paper is intended to be a self-contained user's manual for the PLRP and VDP, and relies on many illustrative examples. I also include quick proofs of two classic results in information economics: I prove the main comparative static for increasing risk aversion in Diamond and Stiglitz (1974) using fewer differentiability assumptions. Pratt's (1988) deduction of the preservation of risk aversion ranking of *interim* utility functions under additional independent risks likewise follows.

Towards a self-contained presentation, I err on the side of including results that overlap with mine, flagging when it happens.[1] To keep track, my novel results are called theorems, gentle consequences of old results are corollaries, and Karlin's VDP is a proposition. Most remaining proofs are in the Appendix.

[1] Most notably, Chade and Swinkels (2019) recently derives a condition very close to my Theorem 1(c). They focus on the LSPM of the signal survivor derivative, with applications primarily in contract theory.

2 Signal Families

2.1 A New Property for General Signals

Consider a kernel $\phi : S \times \Theta \to \mathbb{R}$ on the linearly ordered subsets $\Theta, S \subset \mathbb{R}$ or the integers. We often interpret Θ as states, S as realized signals, and ϕ a transition kernel. The *(strict) monotone likelihood ratio property* holds if for every $s_0 \in S$, the *likelihood ratio* $\ell(s, \theta) \equiv \phi(s, \theta)/\phi(s_0, \theta)$ is (strictly) monotone in θ, for all $s > s_0$. Equivalently, ϕ is (strictly) log-supermodular (LSPM), or $\phi(s_0, \theta_0)\phi(s, \theta) \geq (>)\phi(s_0, \theta)\phi(s, \theta_0)$ if $s > s_0$ and $\theta > \theta_0$. In the differentiable case, this is equivalent to $[\log \phi(s, \theta)]_{s\theta} \geq (>)0$.

We say that ϕ obeys the stronger (strict) *proportionate likelihood ratio property* (PLRP) if not only is the likelihood ratio (increasing) nondecreasing, but its rate of increase is itself log-supermodular. Namely, the (strict) PLRP means the (strict) MLRP, and if for any $\theta_3 > \theta_2 > \theta_1$ and $s > s_0$, the ratio is (increasing) nondecreasing in s:

$$\frac{\ell(s, \theta_3) - \ell(s, \theta_2)}{\ell(s, \theta_2) - \ell(s, \theta_1)} \tag{1}$$

If ϕ is twice differentiable, then it obeys the *smooth proportionate likelihood ratio property* (PLRP*) if the derivative ℓ_θ is log-supermodular for all $\theta \in \Theta$ and $s > s_0$.

Next, Karlin (1968) considers a key property of transition kernels ϕ: In each case below, Φ_n is an arbitrary $n \times n$ matrix with (j, i)-th entry $\phi_{ji} \equiv \phi(s_j, \theta_i)$, where $\theta_1 < \theta_2 < \cdots < \theta_n$ and $s_1 < s_2 < \cdots < s_n$. Then ϕ is *totally positive of order 1* (written, TP-1) if $\phi \geq 0$. Next, ϕ is *totally positive of order 2* (TP-2) if it is TP-1 and the determinant of Φ_2 is non-negative. Recursively, ϕ is *totally positive of order k* (TP-k) if it is TP-$(k-1)$ and the determinant of Φ_k is non-negative. Finally, ϕ is *totally positive* (TP) if it is TP-n for all $n = 1, 2, \ldots$. In particular, TP-1 asserts $\phi \geq 0$, TP-2 asserts $\phi \geq 0$, and ϕ obeys the MLRP. But TP-3 imposes an additional condition whose formulation has escaped simple formulation in the literature. I now address this gap.

Theorem 1 (PLRP) *(a) The kernel ϕ is TP-3 iff it obeys the PLRP.*
(b) If ϕ is twice differentiable in θ, then PLRP and PLRP* are equivalent.
(c) If ϕ is C^3, then (strict) PLRP* holds iff ϕ and $\phi^2 \cdot (\log \phi)_{\theta s}$ are (strictly) LSPM.[2]

[2] I recently learned that this condition was recently derived by Chade and Swinkels (2019) (Lemma 1), using the stronger version of the TP-3 property for C^4 kernels in Karlin (1957) (which does not parse strict and weak cases) — while my proof could be relaxed from C^3 to C^2 (at some increase in complexity). They prove the slightly weaker result that if ϕ obeys the strict MLRP, then strict LSPM of $\phi^2 \cdot (\log \phi)_{\theta s}$ is equivalent to Karlin's sufficient condition for TP-3, and in turn his necessary condition for TP-3 implies weak LSPM of $\phi^2 \cdot (\log \phi)_{\theta s}$.

Often in applications, ϕ is a posterior density in s:

$$\phi(s,\theta) = \frac{f(\theta)h(s|\theta)}{\int f(x)h(s|x)dx} \tag{2}$$

where $f(\theta)$ is the prior density, and $h(s|\theta)$ the conditional signal density. Then ϕ is a *stochastic kernel*, or a density in θ for all s, so with $\phi \geq 0$ and $\int \phi(s,\theta)d\theta = 1$ for all s.

As multiplication by a function of one variable does not affect total positivity,[3] it is well-known that ϕ is TP-n if and only if the conditional signal density h is itself TP-n. So MLRP and PLRP are preserved by multiplication by $f(s), g(\theta) > 0$:

$$\phi(s,\theta) \textit{ is MLRP (PLRP) iff } f(\theta)g(s)\phi(s,\theta) \textit{ is MLRP (PLRP) } \forall f(\theta), g(s) > 0 \quad (\star)$$

Example 1. For an application of Theorem 1, consider a Gamma density with shape parameter α and scale parameter $\theta > 0$, namely, $h(s|\theta) = s^{\alpha-1}e^{-\frac{s}{\theta}}/[\Gamma(\alpha)\theta^{\alpha}]$. Then the stochastic kernel $h(s|\theta)$ obeys both MLRP and PLRP. To see this, write $h(s|\theta) = f(\theta)g(s)e^{-\frac{s}{\theta}}$. Then $h(s|\theta)$ obeys the MLRP and the PLRP iff $\phi(s,\theta) \equiv e^{-\frac{s}{\theta}}$ obeys both. Now, ϕ is LSPM since $(\log\phi)_{s\theta} = 1/\theta^2 > 0$, i.e. obeys the MLRP. Next, by Theorem 1(c), ϕ obeys the PLRP, as $\log(\phi^2 \cdot (\log\phi)_{s\theta}) = 2\log\phi(s,\theta) - 2\log\theta$ is SPM.

Or consider the Gaussian density. For this, given ($\star$), the PLRP depends on the exponential portion $\phi(s,\theta)$, for which $\log\phi(s,\theta) = -(s-\theta)^2/2\sigma^2$. Since $(\log\phi)_{s,\theta} = 1/\sigma^2 > 0$, we see that ϕ is LSPM. Finally, this implies $\phi^2(\log\phi)_{s,\theta} = \phi^2/\sigma^2$ is LSPM.

2.2 A New Property for Location Family Signals

We now assume an additive location signal family, with $S = \Theta \subseteq \mathbb{R}$ and $\phi(s,\theta) = \psi(s-\theta)$, for a twice differentiable function $\psi : \mathbb{R} \to \mathbb{R}^+$. As is well-known, ϕ *obeys the MLRP iff* ψ *is log-concave*, and so iff the *rate of density change* $\delta(s) \equiv (\log\psi(s))'$ is nonincreasing. Theorem 1(c) yields a simple PLRP characterization for location signals.

Corollary 1 *The kernel ϕ obeys the PLRP iff δ and $2\delta + \delta''/\delta'$ are nonincreasing. In particular, if ϕ obeys the MLRP and δ' is log-concave, then ϕ obeys the PLRP.*

Example 2. For the Gaussian density, we see that $\delta(s) = (-s^2/2\sigma^2)' = -s/\sigma^2$. This is decreasing, as is $2\delta + \delta''/\delta' = -2s/\sigma^2$; and so the PLRP* follows from Corollary 1.

[3]TP-n asks that all minor determinants be non-negative in the matrix with (j,i)th entry $\phi(s_j,\theta_i)$. If we multiply ϕ by a function $f : \Theta \to \mathbb{R}^+$, then in the corresponding matrix, each entry in the ith column is scaled by $f(\theta_i)$; this scales each minor determinant involving column i by a factor of $f(\theta_i) \geq 0$, and as such, does not affect total positivity. Similarly for multiplication by a function of s only.

We next apply Corollary 1 in an example that escapes Theorem 1. The extreme value distribution is quite commonly used in economics:[4] $\psi(s) = e^{-s/\sigma}e^{-e^{-s/\sigma}}/\sigma$. The signal $\phi(s,\theta) = \psi(s-\theta)$ thus generated obeys both the MLRP and the PLRP since $\delta(s) = (e^{-s/\sigma}-1)/\sigma$ is decreasing, as is $2\delta(s)+\delta''(s)/\delta'(s) = (2e^{-s/\sigma}-3)/\sigma$.

In many economic applications, such as choice among finitely many options (as often arises in informational herding), the signal is truncated. Log-concavity is an incredibly useful property in such settings, since it is preserved by truncation. Let's review the common reason why: a product of TP-2 functions is itself TP-2, and the indicator function $\mathbb{I}_{\theta\le s}$ is TP-2 (Karlin, 1968), and so the truncated function $\phi(s,\theta)\cdot\mathbb{I}_{\theta\le s}$ is MLRP.

While Karlin (1968) also showed that the indicator function itself is totally positive of all orders n, a product of two TP-3 functions is not generally itself TP-3.[5] In particular, $\phi(s,\theta)\cdot\mathbb{I}_{\theta\le s}$ need not obey PRLP even if $\phi : S\times\theta\to\mathbb{R}$ does. But if the density $\phi(s,\theta)$ vanishes for signals s near θ, then the indicator function preserves the PLRP:

Theorem 2 (Truncation) *Let $\phi : S\times\Theta\to\mathbb{R}$. Then*
(a) $\phi(s,\theta)\mathbb{I}_{\theta\le s}$ obeys the MLRP if $\phi(s,\theta)$ obeys the MLRP.
(b) $\phi(s,\theta)\mathbb{I}_{\theta\le s}$ obeys the PLRP if $\lim_{s\downarrow\theta}\phi(s,\theta)=0$ for all θ, and $\phi(s,\theta)$ obeys the PLRP*.*

Example 3. Consider the Gamma density $\psi(s)\propto\mathbb{I}_{s\ge 0}s^{\alpha-1}e^{-\frac{s}{\beta}}$. Then I claim that the stochastic kernel $h(s|\theta) = \psi(s-\theta)$ obeys the MLRP if $\alpha\ge 1$ and the PLRP if $\alpha\ge 2$. For the rate of change δ *in the untruncated density* obeys:

$$\delta(s) \equiv \left(\log(s^{\alpha-1}e^{-\frac{s}{\beta}})\right)' = \frac{\alpha-1}{s} - \frac{1}{\beta} \quad\Rightarrow\quad 2\delta+\delta''/\delta' = 2\left(\frac{\alpha-2}{s}-\frac{1}{\beta}\right)$$

By Corollary 1, the MLRP holds since δ is decreasing for $\alpha\ge 1$, and the PLRP* since $2\delta+\delta''/\delta'$ is decreasing when $\alpha\ge 2$. Since $\psi(s|\theta)$ is continuous in s, and $\psi(\theta|\theta)=0$, Theorem 2 implies that MLRP and PLRP* survive the indicator function multiplication.

Next, the Chi-squared density $\psi(s)\propto\mathbb{I}_{s\ge 0}s^{k/2-1}e^{-s/2}$ obeys the MLRP if $k\ge 2$, and the PLRP* if $k\ge 4$. For the rate of change δ of the untruncated density:

$$\delta(s) = \left(\left(\frac{k-2}{2}\right)\log s - \frac{s}{2}\right)' = \frac{k-2}{2s} - \frac{1}{2}$$

It suffices to note that this is decreasing, as is the expression $2\delta+\delta''/\delta' = (k-4)/s-1$.

[4]This calculation assumes shape parameter zero. More generally, the distribution obeys the MLRP for any shape parameter in $[-1,0]$ and the PLRP for any shape parameter in $[-0.5,0]$.

[5]Given two TP-3 functions $\phi_1,\phi_2 : S\times\Theta\to\mathbb{R}^+$; it need not be the case that for $\theta_1<\theta_2<\theta_3$ and $s_1<s_2<s_3$, the 3×3 matrix with (j,i)th entry $\phi_1(s_j,\theta_i)\cdot\phi_2(s_j,\theta_i)$ has a non-negative determinant.

The limit condition in Theorem 2(b) is sufficient, but not necessary. For example, assume that ψ is the exponential density. Then $\phi(s,\theta) = \mathbb{I}_{\theta \leq s} \cdot \lambda e^{-\lambda s} e^{\lambda \theta}$. Even though $\phi(\theta,\theta) = \lambda$, this function nonetheless obeys the PLRP, as the product of two functions that each depend on only one variable, and the TP indicator function.

3 Sign Changes of Expected Payoffs

I henceforth assume that $u : \Theta \to \mathbb{R}$ is piecewise continuous and $\phi : S \times \Theta \to \mathbb{R}$ is absolutely continuous w.r.t. a measure μ, with Θ and S subsets of $\mathbb{R}$. I now apply the signal family properties in expectations. For this, I turn to Karlin's *variation diminishing property (VDP)*. Let us call $f : \mathbb{R} \mapsto \mathbb{R}$ *n-crossing* if n is the maximum number of strict sign changes, $\{+\}$ to $\{-\}$, or vice versa, across all possible increasing sequences (t_i), namely:

$$\{(f(t_1), \ldots, f(t_k)), t_1 < \cdots < t_k, \forall k \in \mathbb{N}\}$$

Say that f has *n weak sign changes* if this is the maximum number of swaps from $\{0,+\}$ to $\{0,-\}$, or vice versa, in any ordered sequence. Karlin derived total positivity conditions on transition kernels (namely, $\int \phi(s,\theta) d\mu(\theta) \equiv 1$ is not needed) ϕ ensuring that the expectation

$$U(s) = \int \phi(s,\theta) u(\theta) d\mu(\theta)$$

has at most as many sign crossings as $u(\theta)$, and in the same order, i.e. first minus to plus, or vice versa. Call u *nondegenerate* if $u \neq 0$ on a positive-measure subset of Θ.

Proposition 1 (VDP) *Let u be n-crossing, initially $+$ to $-$ ($-$ to $+$).*
(a) [WEAK] *If ϕ is TP-$(n+1)$, then $U(s)$ cannot have more than n sign changes, and is initially $+$ ($-$) if it is exactly n-crossing.*
(b) [STRICT] *If ϕ is* strictly *TP-$(n+1)$, and u is nondegenerate, $U(s)$ cannot have more than n weak sign changes, and is initially* $\{0,+\}$ ($\{0,-\}$) *if it has n weak sign changes.*

In particular, if ϕ obeys the MLRP (and so is TP-2), then the VDP says that if a function u is upcrossing[6] in θ, then $U(s)$ is upcrossing in s.[7] And if ϕ is PLRP (and so is TP-3), then the expectation of a two-crossing function is at most two-crossing, and in the same sign change order. By the strict VDP, U cannot have any zero subintervals, as this would constitute infinitely many weak sign changes.

[6]A function is *upcrossing* (respectively, *downcrossing*) if it crosses the horizontal axis at most once, and if once, from below (respectively, above).

[7]This important result was first proved in Karlin and Rubin (1956).

EXAMPLE 4. Assume a utility function $u(x, y)$ for wealth x, whose Arrow-Pratt risk aversion coefficient $r_y(x) \equiv -u_{xx}/u_x$ falls in a parameter y. Namely, the marginal utility of money $u_x(x, y)$ is LSPM. Diamond and Stiglitz (1974) find that the risk premium rises in risk aversion, or equivalently, the certainty equivalent rises in y. This follows if $E_X u(X, y) - u(z, y)$ is upcrossing in y, for any z. For then $E_X u(X, y) = u(z, y)$ implies $E_X u(X, y') > u(z, y')$ for all $y' > y$, and therefore the certainty equivalent at y' exceeds z. Finally, if X has a density f and cdf F over wealth, integration by parts gives[8]

$$\int_{\underline{x}}^{\bar{x}} u(x, y) f(x) dx - u(z, y) = \int_{\underline{x}}^{\bar{x}} u_x(x, y)(\mathbb{I}_{x \geq z} - F(x)) dx \tag{3}$$

Since $\mathbb{I}_{x \geq z} - F(x)$ is upcrossing in x, expectation (3) is upcrossing in y, by Proposition 1.

Karlin (1968) offers a simple application of the VDP for stochastic kernels:[9]

Corollary 2 (Preservation) *Let (ϕ, μ) define a stochastic kernel, and $u : \Theta \to \mathbb{R}$.*
(a) If ϕ obeys the MLRP, and u is monotone up (down), then so is $U(s)$.
(b) If ϕ obeys the PLRP, and u is quasiconcave (quasiconvex), then so is $U(s)$.
(c) If premise (a) or (b) is strict, then $U(s)$ nowhere constant[10] if $u(\theta)$ is not constant.

Proof: Any function u is monotone up (down) iff $u(\theta) - c$ is upcrossing (downcrossing), for all $c \in \mathbb{R}$. If ϕ obeys the MLRP, then $\int \phi(s, \theta)(u(\theta) - c) d\mu(\theta) = U(s) - c$ is upcrossing (downcrossing) for all $c \in \mathbb{R}$, by Proposition 1. So $U(s)$ is monotone up (down). A function $u(\theta)$ is quasiconcave if it crosses any horizontal line at most twice, initially from below if twice. The proof now proceeds identically using the PLRP.

Part (c) follows from the strict VDP (Proposition 1). □

The famous interpretation in Milgrom (1981) of Corollary 2(a) was that an MLRP kernel ϕ ensures the *favorable news* rank order: Specifically, higher signals $s_2 > s_1$ lead to higher expected payoffs $U(s_2) \geq U(s_1)$ for any monotone payoff function $u(\theta)$. Say that signals have the *harmonious news property* if whenever $u(\theta)$ is quasi-concave (convex), then so too $U(s)$. So just as the MLRP ensures the favorable news property, PLRP guarantees the harmonious news property, by Corollary 2(b).

EXAMPLE 5. In Example 4, assume final wealth X is the sum of two independent r.v.s S and W, with S drawn with density f. Pratt (1988) sought conditions ensuring that

[8] Diamond and Stiglitz (1974) assumed three derivatives; this logic requires one.
[9] He omits Corollary 2(b), but it's an easy consequence.
[10] Namely, there is no open subinterval of $\mathbb{R}$ on which $U(s)$ is constant.

the risk aversion ranking of *interim* utility functions $U(w,y) = \int f(s)u(w+s,y)ds = \int f(x-w)u(x,y)dx$ agrees with the ex post utility risk aversion ranking of $u(x,y)$.

Pratt showed that the interim risk aversion coefficients $R_y(w) \equiv -U_{ww}/U_w$ are ranked $R_{y_1}(w) \geq R_{y_2}(w)$ when $y_1 < y_2$ if $r_{y_1}(x) \geq r(x) \geq r_{y_2}(x)$ for all x, for some decreasing function $r(x)$. To quickly secure his result here, integrate both the numerator and denominator of $R_y(w) \equiv -U_{ww}/U_w$ by parts[11] to get

$$R_y(w) = \int \phi(x,y|w)r(x,y)dx \qquad \text{where} \quad \phi(x,y|w) \equiv \frac{f(x-w)u_x(x,y)}{\int f(x-w)u_x(x,y)dx} \tag{4}$$

Now, u_x is LSPM since $-u_{xx}/u_x$ falls in y, and thus the stochastic kernel $\phi(x,y|w)$ is LSPM in (x,y). Since $r(x)$ is decreasing, its monotonicity is preserved by Corollary 2(a), and

$$R_{y_1}(w) \geq \int \phi(x,y_1|w)r(x)dx \geq \int \phi(x,y_2|w)r(x)dx \geq R_{y_2}(w)$$

Example 6. To see that the MLRP need not preserve quasiconcavity, consider the kernel matrix:

$$\Phi = [\phi(s_j,\theta_i)] = \begin{pmatrix} 9 & 9 & 0 \\ 3 & 6 & 9 \\ 2 & 6 & 10 \end{pmatrix}$$

All 2x2 minors of Φ are positive, and therefore Φ is TP-2, and thus ϕ obeys the MLRP. But the determinant of Φ is -54, and thus ϕ is not TP-3, and so violates the PLRP. Finally, for the quasiconcave function $u(\cdot)$ given by $u(\theta_1) = -3, u(\theta_2) = 4, u(\theta_3) = -1$, the expectation $\Phi u = (9,6,8)$ is strictly quasi*convex*, and so not quasiconcave.

This result speaks to the key difference between the economic applications for the MLRP and the PLRP. The MLRP applies to "vertical" preference models, with monotone preferences. For example, if a firm earns a profit $u(\theta) = \theta$ from hiring an employee of type θ, then its *expected* profit $\Pi(s)$ is also monotone in the signal realization s, by Corollary 2(a). On the other hand, many economic settings assume a "variety" rather than "quality" payoff structure. The PLRP is designed for such "horizontal" matching environments. Voters, e.g., prefer political candidates closer to them. Also, in matching models, individuals frequently seek out more kindred partners. To be specific, if a firm of type τ incurs the quadratic loss $(\tau-\theta)^2$ from hiring an employee of type θ, then its *expected* payoff $\Pi(s)$ is also quasiconcave in the signal realization s, by Corollary 2(b).

Let cdf $\Phi(s,\theta) = \int_{-\infty}^{\theta} \phi(s,t)d\mu(t)$ correspond to stochastic kernel $\phi(s,\theta)d\mu(\theta)$.

[11] For $U_w = \int -f'(x-w)u(x,y)dx = \int f(x-w)u_x(x,y)dx$, and $-U_{ww} = \int f(x-w)u_x(x,y)r(x,y)dx$.

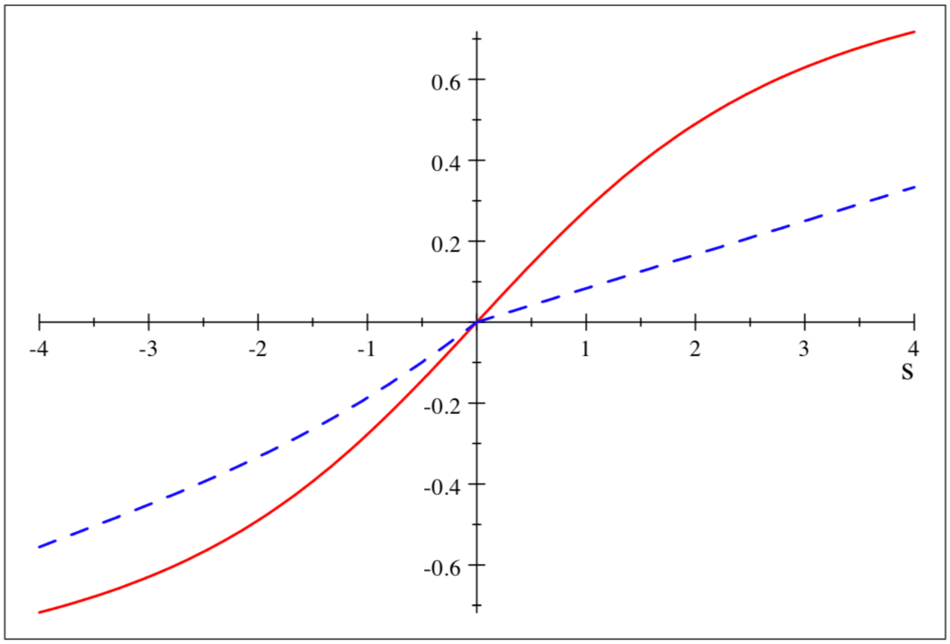

Figure 1: **Vertical Example.** Assume a uniform prior on $\Theta = [-1, 1]$. We consider the MLRP kernels $\phi_1(s, \theta) = c_1(s)e^{-\frac{1}{2}(s-\theta)^2}$ and $\phi_2(s, \theta) = c_2(s)(\theta s + 4)\mathbb{I}_{\theta \leq 1+s/4}$. The constant $c_i(s)$ ensures that ϕ_i is a probability density on $\theta \in [-1, 1]$ for each s. For the monotone function $u(\theta) = \theta$, each expected payoff $\bar{\theta}_i(s) \equiv \int_{-1}^{1} \phi_i(s, \theta)\theta d\theta$ is monotone — solid red and dashed blue respectively for $i = 1, 2$.

Corollary 3 (Stochastic Shifts) *Let $(\phi, \mu) : S \times \Theta \to \mathbb{R}_+$ define a stochastic kernel.*
(a) If ϕ obeys the (strict) MLRP, then $\Phi(s_2, \theta) \leq (<)\Phi(s_1, \theta)$ for all θ and $s_2 > s_1$
(b) If ϕ obeys the (strict) PLRP, and Φ is differentiable in s, then $-\Phi_s(s, \theta)$ obeys the (strict) MLRP.[12]

Part (a) follows from Corollary 2(a) and (c), writing $\Phi(s, \theta) = \int \phi(s, t)u(t)d\mu(t)$ with $u(t) = \mathbb{I}_{t \leq \theta}$ —- which is obviously nonincreasing in t for any θ.

Corollary 3(b) offers another way to prove Corollary 2(b): Integrating by parts, $\Pi'(s) \equiv \int -\Phi_s(s, \theta) \cdot 2(\tau - \theta)d\theta$ is downcrossing by Proposition 1, i.e. Π is quasi-concave.

Example 7. To see that the MLRP alone cannot guarantee that $-\Phi_s(s, \theta)$ is LSPM, consider the stochastic kernel $\phi(s, \theta) = \frac{1}{8}(\theta s + 4)$ for $s \in [-4, 4]$ with support $\theta \in [-1, 1]$.

[12]The weak part of this result is implied by Proposition 1(ii) and Lemma 1 in Chade and Swinkels (2019). Their proof is direct and algebraic. I give an indirect (but simpler) proof using the VDP, showing that if $-\Phi_s$ were not LSPM, then one could construct a contradiction to Corollary 2(b).

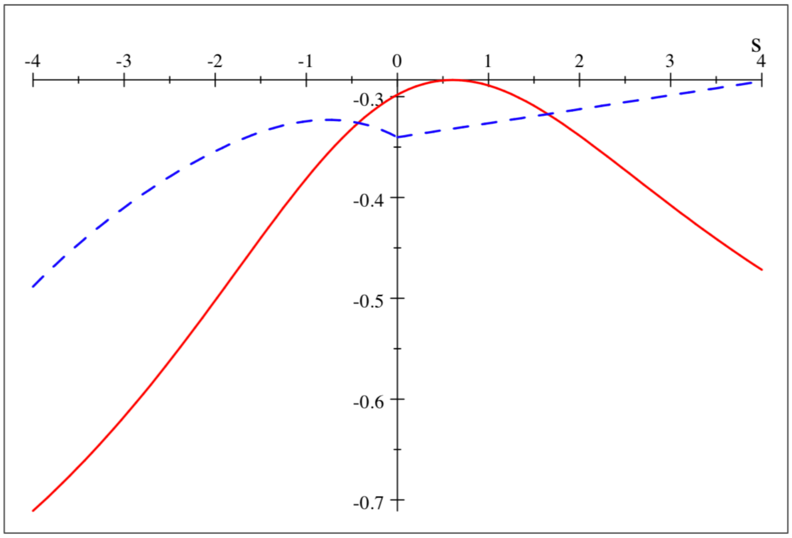

Figure 2: **Horizontal Example.** For the stochastic kernels in Figure 1, I plot the expectations $\Pi_i(s) \equiv \int_{-1}^{1} \phi_i(s,\theta)u(\theta)d\theta$ of the quasiconcave function $u(\theta) = -(\theta - 1/12)^2$ (so red and blue are $i = 1, 2$, respectively). The PLRP kernel ϕ_1 preserves quasiconcavity, but ϕ_2 is not a PLRP kernel (by footnote 13), and does not preserve quasiconcavity. Finally, this plot shows that PLRP does not preserve concavity.

This obeys the strict MLRP, with

$$(\log \phi)_{s\theta} = \frac{4}{(s\theta + 4)^2} = \frac{1}{16\phi^2} > 0 \tag{5}$$

But ϕ is not strictly PLRP, for then $\phi^2(\log \phi)_{s\theta}$ would be log-supermodular, by Theorem 1(c). But in fact, $16\phi^2(\log \phi)_{s\theta} = 1$ is constant, from (5). Finally, check that

$$\Phi(s,\theta) \equiv \int_{-1}^{\theta} \tfrac{1}{8}(ts+4)dt = \tfrac{1}{16}(1+\theta)(s(\theta-1)+8) \quad \Rightarrow \quad -\Phi_s(s,\theta) = \tfrac{1}{16}(1-\theta^2)$$

Then $-\Phi_s(s,\theta)$ is not strictly MLRP, since it is constant in s.[13]

Since the PLRP preserves quasiconcavity and quasiconvexity, one might wonder if it does the same for concavity and convexity. This is false. In particular, the solid red curve in Figure 2 becomes strictly convex near the edges of the region shown. When the posterior mean $\overline{\theta}(s) \equiv \int \phi(s,\theta)\theta d\theta$ is concave (convex), and ϕ is a PLRP kernel, then provided $u(\theta)$ is monotone and concave (convex), so too is the expectation $U(s)$. More generally, any PLRP kernel ϕ amplifies concavity and convexity:

[13] The stochastic kernel ϕ was generated using a uniform prior on $\Theta = [-1,1]$ and the signal density $h(s|\theta) = (\theta s + 4)$. But scaling the density with indicator $\mathbb{I}_{\theta \le 1+s/4}$, as in Figure 2, results in a kernel ϕ that violates even the weak PLRP, and in fact $-\Phi_s$ violates the MLRP.

Theorem 3 (Amplification) *Assume that the stochastic kernel* (ϕ, μ) *obeys the PLRP, or that* $-\Phi_s$ *obeys the MLRP. If* $u(\theta)$ *is a concave (convex) function of an increasing function* $v(\theta)$*, then* $U(s)$ *is a concave (convex) function of* $V(s) \equiv \int \phi(s,\theta)v(\theta)d\mu(\theta)$.

The proof in A.8 relies only on the MLRP property $-\Phi_s$. For a stochastic kernel of a location family $\phi(s,\theta)d\mu(\theta) \equiv \psi(s-\theta)d\theta$, so that $-\Phi_s(s) = \int_{-\infty}^{\theta} -\psi'(s-t)dt = \phi(s,\theta)$, the result therefore only requires that ϕ obey the MLRP. If ϕ obeys the stronger PLRP, Theorem 3 follows from Proposition 1, as noted in Example 6 in Jewitt (1987).

Example 8. Revisiting Example 5, Theorem 3(b) offers another condition for the interim utility function $U(w) = E[u(w+S)] = \int f(x-w)u(x)dx$ to preserve the risk aversion order of u. If u is more concave than v, then U is more concave than $E[v(w+S)]$ provided f obeys the PLRP. And since f is a location density, it suffices that f obey the MLRP, by the remark after Theorem 3. In particular, following Pratt (1988), the risk premium rises in risk aversion when final wealth $X = S + W$ is the sum of two independent random variables, provided that one of them has a log-concave density.[14]

Example 9. In Holmstrom (1979), a principal chooses a profit-sharing rule $a : \Theta \to \mathbb{R}$ to maximize $\int \phi(s,\theta)(\theta - a(\theta))d\theta$, where the density ϕ over profits θ depends on the agent's (unobserved) effort s. A risk-averse agent chooses s to maximize expected utility of his profit share less effort costs, namely $U(s) \equiv \int \phi(s,\theta)u(a(\theta))d\theta - c(s)$. The standard first-order approach replaces the agent's incentive constraint with the F.O.C. $U'(s) = 0$. Jewitt (1988) then sought conditions justify the first-order approach, observing that if effort costs are convex, then it suffices that $\int \phi(s,\theta)u(a(\theta))d\theta$ is concave in s. And by Theorem 3, this holds provided that (i) ϕ obeys the PLRP, (ii) expected profits $\int \phi(s,\theta)\theta d\theta$ are weakly concave in effort s, and (iii) $u(a(\theta))$ is concave in θ.

4 Sign Changes of Optimal Actions

Consider a decision maker DM who earns a differentiable payoff $u(a,\theta)$ from choosing action $a \in A$ in state $\theta \in \Theta$. I now assume that $A, \Theta \subset \mathbb{R}$ are compact subsets. Milgrom and Shannon (1994) established that the optimal action rule is monotone in the strong set order provided $u(a',\theta) - u(a,\theta)$ is strictly upcrossing[15] in θ, for any $a' > a$. If u is differentiable in a, then it suffices that marginal utility $u_a(a,\theta)$ is weakly increasing in

[14] This is closely related to Theorem 5 in Jewitt (1987).

[15] $f : \mathbb{R} \mapsto \mathbb{R}$ is *strictly upcrossing* in x if whenever f is nonnegative (positive) at x, it is nonnegative (positive) at any $x' > x$. Weakly upcrossing allows weakly positive to follow strictly positive.

θ for all a. And if u_a is strictly increasing in θ, or if u is also strictly concave in a, then the optimal action is monotone — *any* selection of optimal actions is monotone in θ.[16]

Assume that the DM sees only a signal $s \in S$ about θ, yielding a posterior density $\phi(\cdot|s)$. Let $V(a,s) = \int_{\theta\in\Theta} \phi(\theta|s)u(a,\theta)d\theta$ be the expected utility of action a given s, and define $a^*(s) = \arg\max_{a\in A} V(a,s)$. If the kernel ϕ obeys the MLRP, and $u_a(a,\theta)$ is monotone up (down) in θ, then $a^*(s)$ is monotone up (down) in the strong set order (SSO). For given $u_a(a,\theta)$ monotone in θ, we have $V_a(a,s)$ monotone in s, by Corollary 2(a). But then $a^*(s)$ is monotone in the SSO, by Milgrom and Shannon (1994).[17] In fact, any selection from optimal action map $a^*(s)$ is monotone if V_a is strictly increasing in s.[18]

I now ask when the argmax $a^*(s)$ is quasiconcave or *quasiconcave in the SSO*, so either monotone up, or monotone down, or first monotone up and then monotone down.

Corollary 4 *Let $a^*(s) = \arg\max_{a\in A} \int_{\theta\in\Theta} \phi(\theta|s)u(a,\theta)d\theta$. If ϕ obeys the PLRP and $u_a(a,\theta)$ is hump-shaped (U-shaped) in θ, then $a^*(s)$ is quasiconcave (quasiconvex) in the SSO. If u is also strictly concave in a, or if ϕ obeys the strict PLRP and $u_{a\theta}$ is strictly downcrossing (upcrossing) in θ, then $a^*(s)$ is quasiconcave (quasiconvex) in s.*

Proof: Corollary 2(b) gives the first claim: For quasiconcavity of $u_a(a,\theta)$ in θ yields quasiconcavity of $V_a(a,s)$ in s, and thus V_{as} is downcrossing. As above, $V_{as} \geq 0$ ($V_{as} \leq 0$) implies an optimal action that is monotone up (down) in the SSO. For the final assertion, if ϕ obeys the strict PLRP and $u_{a\theta}$ is strictly downcrossing, then by the strict VDP (Proposition 1), V_a is quasiconcave in s, and level on no subintervals. Thus, V_{as} is downcrossing and can only vanish at a single point s. When $V_{as} > 0$, the optimal action is monotone up; when $V_{as} < 0$, the optimal action is monotone down. □

Example 10. Assume $u(a,\theta) = -(a-\theta)^2$. Since $u(a,\theta)$ is strictly concave in a, so is $V(a,s) = \int_{\theta\in\Theta} \phi(\theta|s)u(a,\theta)d\theta$, and there is a unique optimal action $a^*(s)$. And since $u_a = -2a + 2\theta$ is increasing in θ, Corollary 4 gives a monotone optimal action. In fact, $a^*(s) = \bar{\theta}_1(s)$ — the increasing function in Figure 1 — via the FOC:

$$\frac{d}{da}\frac{-\int_{-1}^{1} e^{-\frac{1}{2}(s-\theta)^2}(a-\theta)^2 d\theta}{\int_{-1}^{1} e^{-\frac{1}{2}(s-\theta)^2} d\theta} = 0 \quad \Rightarrow \quad a^*(s) = \frac{\int_{-1}^{1} e^{-\frac{1}{2}(s-\theta)^2}\theta d\theta}{\int_{-1}^{1} e^{-\frac{1}{2}(s-\theta)^2} d\theta} \equiv \bar{\theta}_1(s)$$

[16] For if $u_{a\theta} > 0$, then $a_2 > a_1$ and $\theta' > \theta$ imply $\int_{a_1}^{a_2}(u_a(a,\theta') - u_a(a,\theta))da > 0$. So if a_2 is optimal at θ, i.e. $\int_{a_1}^{a_2} u_a(a,\theta)da \geq 0$, then $a_1 < a_2$ is suboptimal at θ'. If u is strictly concave in a, then uniqueness of the optimal action implies that monotonicity in the SSO coincides with standard monotonicity.

[17] It is a monotone function — not just a monotone SSO correspondence — given a unique optimal action; this happens if u (and so V) is strictly concave in a.

[18] By Corollary 2(c), this premise follows if ϕ obeys the strict MLRP and u_a is strictly monotone in θ.

Likewise, $u(a,\theta) = -(a-\theta^2)^2$ has a unique optimal choice $a^*(s)$. Since $u_a(a,\theta) = -2(a-\theta^2)$ is U-shaped in θ, the optimal action $a^*(s)$ is quasiconvex in s, by Corollary 4.

Example 11. Bob believes that his effort b is a complement to his co-author Ann's productivity θ for $\theta < 1$, but becomes a substitute once $\theta > 1$. Given effort cost $\frac{1}{2}b^2$, Bob's payoff is

$$u^B(b,\theta) = \begin{cases} b\theta - \frac{1}{2}b^2 & \text{if } \theta \in [0,1] \\ b(2-\theta) - \frac{1}{2}b^2 & \text{if } \theta \in (1,2] \end{cases}$$

But Bob only sees Ann's effort a, and thinks that θ has a density $\phi(\theta|a)$ that obeys the PLRP. By Corollary 4, Bob's optimal effort is quasiconcave in Ann's observable effort.

5 A Peasant's Proof of the VDP

The proof of Karlin's variation diminishing property for single-crossing functions with TP-2 kernels is straightforward. But Karlin's proof of the VDP, Proposition 1, is difficult to follow, since it uses Cauchy-Binet matrix decompositions that are not standardly taught in linear algebra. This may explain why it has seldom been applied in economics.

I now offer a novel "peasant's proof" of Karlin's VDP within reach of freshman linear algebra. To understand it, notice that the contrapositive of the VDP would assert that (a) if $U(s) = \int \phi(s,\theta)u(\theta)d\mu(\theta)$ crosses zero more than n times, then $u(\theta)$ must cross more than n times; and (b) if $U(s)$ is n-crossing and initially downcrossing, then $u(\theta)$ is not n-crossing and initially upcrossing. I prove an equivalent contrapositive result:

Lemma 1 (VDP Contrapositive) *Let $\Theta \subset \mathbb{R}$ be finite or a subinterval. Assume $\int \phi(s,\theta)u(\theta)d\mu(\theta)$ is n-crossing (has n weak sign changes) on some $S' \subseteq S$, initially $-$ to $+$ ($+$ to $-$), and ϕ is TP-$(n+1)$ (resp. ϕ is strictly TP-$(n+1)$ and u nondegenerate). Then u is n-crossing, initially $-$ to $+$ ($+$ to $-$) on some subset $\Theta' \subseteq \Theta$.*

Proof: First, posit a finite state space $\Theta = \{\theta_i | \theta_1 < \cdots < \theta_N\}$. I use induction on $n < N$.

The Case $n = 1$. Let ϕ be TP-2. Picking a case, assume $\sum_{\theta\in\Theta} \phi(s,\theta)u(\theta)$ crosses $+$ to $-$ on $S' \subseteq S$. Pick $s_1 < s_2$ such that (‡) $\alpha_1 \equiv \sum_{\theta\in\Theta} \phi(s_1,\theta)u(\theta) \geq 0$ while $\alpha_2 \equiv \sum_{\theta\in\Theta} \phi(s_2,\theta)u(\theta) \leq 0$, each inequality strict for the weak VDP. To prove that u crosses $+$ to $-$ on some $\Theta' \subseteq \Theta$, it suffices to prove that it is not upcrossing on Θ.

For a contradiction, pick $\theta^* \in \Theta$ with $u(\theta) \leq 0$ for $\theta \leq \theta^*$ and $u(\theta) \geq 0$ for $\theta > \theta^*$. Then:

$$\alpha_2 \equiv \sum_{\theta\leq\theta^*} \frac{\phi(s_2,\theta)}{\phi(s_1,\theta)}\phi(s_1,\theta)u(\theta) + \sum_{\theta>\theta^*} \frac{\phi(s_2,\theta)}{\phi(s_1,\theta)}\phi(s_1,\theta)u(\theta)$$

Given TP-2, $\phi(s_2,\theta)/\phi(s_1,\theta)$ rises in θ, and $u(\theta) \leq 0$ in the first sum and $u(\theta) \geq 0$ in the second, we have a contradiction to $\alpha_2 < 0$:

$$\alpha_2 \geq \frac{\phi(s_2,\theta^*)}{\phi(s_1,\theta^*)} \sum_{\theta \leq \theta^*} \phi(s_1,\theta)u(\theta) + \frac{\phi(s_2,\theta^*)}{\phi(s_1,\theta^*)} \sum_{\theta > \theta^*} \phi(s_1,\theta)u(\theta) = \frac{\phi(s_2,\theta^*)}{\phi(s_1,\theta^*)}\alpha_1 \geq 0 \qquad (6)$$

For the strict VDP, strict TP-2 implies a strict first inequality in (6) (i.e. not $\alpha_2 \leq 0$) unless $\phi(s_1,\theta)u(\theta) = 0$ for all $\theta \neq \theta^*$. If this holds then (‡) reduces to $\alpha_1 = \phi(s_1,\theta^*)u(\theta^*)$, which is strictly negative by nondegeneracy and $u(\theta^*) \leq 0$. This contradicts $\alpha_1 \geq 0$.

Inductive Hypothesis. Assume that if $\sum_\theta \phi(s,\theta)u(\theta)$ is $(n-1)$-crossing (has n weak sign changes for the strict VDP) with an initial $+$ to $-$ ($-$ to $+$) on $S' \subseteq S$, and ϕ is TP-n (resp. strictly TP-n with u nondegenerate), then $u(\cdot)$ is $(n-1)$-crossing with an initial downcrossing (upcrossing) on some $\Theta' \subseteq \Theta$.

Proof for n Crossings: Assume ϕ is TP-$(n+1)$, and that $\sum_{\theta\in\Theta} \phi(s_j,\theta)u(\theta)$ is n-crossing with an initial downcrossing on $S' \subseteq S$. So we can choose $s_1 < \cdots < s_{n+1}$ in S, and reals $\{\alpha_j\}_{j=1}^{n+1}$ with alternating signs $(-1)^{j+1}\alpha_j > 0$ (and ≥ 0 for the strict VDP), with:

$$\textstyle\sum_{i=1}^N \phi(s_j,\theta_i)u(\theta_i) = \alpha_j \qquad \forall j \in \{1,2,\ldots,n+1\} \qquad (7)$$

We need to prove that $u(\theta)$ is n-crossing on some $\Theta' \subseteq \Theta$ with an initial $+$ to $-$. Now, by (7), $\sum_{\theta\in\Theta} \phi(s_j,\theta)u(\theta)$ is $(n-1)$-crossing (has n weak sign changes, for the strict VDP) on $\{s_1,\ldots,s_n\}$ with an initial downcrossing, and similarly $(n-1)$-crossing (has n weak sign changes, for the strict VDP) on $\{s_2,\ldots,s_{n+1}\}$ with an initial upcrossing. By the inductive hypothesis, $u(\theta)$ is $(n-1)$-crossing with an initial $+$ to $-$ on some $\Theta_1 \subseteq \Theta$, and $(n-1)$-crossing with an initial $-$ to $+$ on some $\Theta_2 \subseteq \Theta$, and thus has at least n total crossings. We need only rule out that $u(\theta)$ is n-crossing on Θ, with an initial $-$ to $+$.

For a contradiction, let $u(\theta)$ be n-crossing on Θ with an initial $-$ to $+$. So we can choose $i_1 < i_2 < \cdots < i_n$ such that $u(\theta_i) \leq 0$ for $i < i_1$, $u(\theta_i) \geq 0$ for $i \in (i_1,i_2)$, and alternating so on, upcrossing at i_n iff n is odd.[19] Choose an $i^* > i_n$ with $u(\theta_{i^*}) \neq 0$, so that

$$u(\theta_i^*) > 0 \text{ if } n \text{ is odd}, \; u(\theta_i^*) < 0 \text{ if } n \text{ is even} \qquad (8)$$

We have identified an index set $\mathcal{I} \equiv \{i^*, i_1, \ldots, i_n\}$. Since (7) consists of $n+1$ equations

[19]Naturally, (i_1,i_2) is the set of indices i_k with $i_1 < i_k < i_2$.

in $N \geq n+1$ unknowns, I will only solve for the unknowns $u(\theta_i)$ for $i \in \mathcal{I}$, namely:

$$\begin{pmatrix} \phi_{1,i^*} & \phi_{1,i_1} & \cdots & \phi_{1,i_n} \\ \phi_{2,i^*} & \phi_{2,i_1} & \cdots & \phi_{2,i_n} \\ \vdots & \vdots & & \vdots \\ \phi_{n+1,i^*} & \phi_{n+1,i_1} & \cdots & \phi_{n+1,i_n} \end{pmatrix} \begin{pmatrix} u(\theta_{i^*}) \\ u(\theta_{i_1}) \\ \vdots \\ u(\theta_{i_n}) \end{pmatrix} = \begin{pmatrix} \alpha_1 - \sum_{i \notin \{i^*, i_1, \ldots, i_n\}} \phi_{1,i} u(\theta_i) \\ \alpha_2 - \sum_{i \notin \{i^*, i_1, \ldots, i_n\}} \phi_{2,i} u(\theta_i) \\ \vdots \\ \alpha_{n+1} - \sum_{i \notin \{i^*, i_1, \ldots, i_n\}} \phi_{n+1,i} u(\theta_i) \end{pmatrix}$$

where I abbreviate $\phi_{j,i} \equiv \phi(s_j, \theta_i)$. Define the matrix $\Phi \equiv [\phi_{j,i}]$ with $j \in \mathcal{J} \equiv \{1, 2, \ldots, n+1\}$ and $i \in \mathcal{I}$. Here, I simply assume Φ non-singular. Appendix §A.4 explores the singular case, which is possible if ϕ is weakly but not strictly TP-$(n+1)$.

Let $\delta(i, i_1, \ldots, i_n)$ and $\beta(i_1, \ldots, i_n)$ be the determinants of Φ, but with the first column replaced by $(\phi_{j,i})_{j \in \mathcal{J}}$ and $(\alpha_j)_{j \in \mathcal{J}}$, respectively. As a determinant is linear in the first column, Cramer's rule yields:

$$u(\theta_{i^*}) = \frac{\beta(i_1, i_2, \ldots, i_n) - \sum_{i \notin \{i^*, i_1, i_2, \ldots, i_n\}} \delta(i, i_1, i_2, \ldots, i_n) u(\theta_i)}{\delta(i^*, i_1, i_2, \ldots, i_n)} \tag{9}$$

I prove that $u(\theta_i^*) \leq 0$ if n is odd, and $u(\theta_i^*) \geq 0$ if n is even, contradicting (8). This follows if $\beta(i_1, \ldots, i_n) \geq 0$, with all $\delta(i, i_1, i_2, \cdots, i_n) u(\theta_i) \leq 0$, and denominator $\delta(i^*, i_1, i_2, \cdots, i_n) < 0$ (> 0) if n is odd (even).

Since the kernel ϕ is TP-$(n+1)$, any matrix determinant $\delta(i, i_1, i_2, \ldots, i_n)$ is nonnegative if $i < i_1$ (where $u(\theta_i) \leq 0$, by construction), nonpositive if $i \in (i_1, i_2)$ (where $u(\theta_i) \geq 0$), etc.[20] In each case, $\delta(i, i_1, i_2, \ldots, i_n) u(\theta_i) \leq 0$. Similarly, the determinant $\beta(i_1, \ldots, i_n) \geq 0$, as $(-1)^{j+1} \alpha_j \geq 0$. All told, the summation in the numerator of (9) is nonnegative, and the denominator is negative iff n is odd, contradicting (8).

Finally, if $\Theta \subset \mathbb{R}$, the proof goes through directly except (7). To get this equation, partition Θ into N equally sized intervals (I_i). Let θ_i be the midpoint of I_i, and define $\phi_{j,i} \equiv \phi(s_j, \theta_i) \mu(I_i)$, where μ is the prior measure over Θ. Choose a partition fine enough that the sequence $(u(\theta_i))_{i=1}^N$ is n-crossing with an initial upcrossing, and such that $\int \phi(s_j, \theta) u(\theta) d\mu(\theta)$ has the same sign as $\alpha_j \equiv \sum_{i=1}^N \phi_{j,i} u(\theta_i)$ for $j \in \{1, 2, \ldots, n+1\}$. Such a selection is possible by the continuity of the functions ϕ and u. □

[20] If $i \in (i_m, i_{m+1})$, then rearranging the columns in ascending order — so as to obtain a matrix with a non-negative determinant, by TP-$(n+1)$ — entails swapping column i with i_1, then i with i_2,..., then i with i_m; a total of m adjacent column swaps, which scales the determinant by a factor of $(-1)^m$.

6 Conclusion

The VDP is an important and increasingly applied result in optimization theory. This paper makes two contributions on this tool in monotone comparative statics under uncertainty. First, since Karlin (1968) text is an important text that is somewhat opaque, and long out of print, I offer a novel and simple proof of the VDP.[21] I largely flesh out the theory of the proportionate likelihood ratio property — a tractable formulation of Karlin's total positivity of order 3 that has only seen sporadic application in economics.[22]

A Omitted Proofs

A.1 Equivalence of TP-3 and PLRP: Proof of Theorem 1 (a)

$\phi : S \times \Theta \to \mathbb{R}$ is TP-3 iff it is TP-2, and also for any $\theta_1 < \theta_2 < \theta_3$ and $s_1 < s_2 < s_3$, with $\phi_{ji} \equiv \phi(s_j, \theta_i)$, the 3×3 determinant of (ϕ_{ji}) must be positive. Expanding it yields:

$$\begin{aligned} & \phi_{31} \begin{vmatrix} \phi_{12} & \phi_{13} \\ \phi_{22} & \phi_{23} \end{vmatrix} - \phi_{32} \begin{vmatrix} \phi_{11} & \phi_{13} \\ \phi_{21} & \phi_{23} \end{vmatrix} + \phi_{33} \begin{vmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{vmatrix} \\ = \; & \phi_{11}\phi_{12}\phi_{13} \left[\left(\frac{\phi_{33}}{\phi_{13}} - \frac{\phi_{32}}{\phi_{12}} \right) \left(\frac{\phi_{22}}{\phi_{12}} - \frac{\phi_{21}}{\phi_{11}} \right) - \left(\frac{\phi_{32}}{\phi_{12}} - \frac{\phi_{31}}{\phi_{11}} \right) \left(\frac{\phi_{23}}{\phi_{13}} - \frac{\phi_{22}}{\phi_{12}} \right) \right] \end{aligned}$$

So ϕ is (strictly) TP-3 iff, in addition to being (strictly) TP-2, the square bracketed term is nonnegative (positive). Defining $\ell(s,\theta) = \phi(s,\theta)/\phi(s_1,\theta)$ for $s > s_1$, write this as

$$\frac{\ell(s_3,\theta_3) - \ell(s_3,\theta_2)}{\ell(s_3,\theta_2) - \ell(s_3,\theta_1)} \geq (>) \frac{\ell(s_2,\theta_3) - \ell(s_2,\theta_2)}{\ell(s_2,\theta_2) - \ell(s_2,\theta_1)}$$

Since this holds for any $s_3 > s_2$, this rearranges to the PLRP Condition:

$$\frac{\ell(s,\theta_3) - \ell\,(s,\theta_2)}{\ell(s,\theta_2) - \ell(s,\theta_1)} \uparrow \text{ in } s \tag{10}$$

A.2 Equivalence of PLRP and PLRP*: Proof of Theorem 1(b)

We show that (10) holds iff ℓ_θ is LSPM, assuming that ℓ_θ exists and $\ell_\theta \geq 0$ (for TP-2).

[21] Park and Smith (2008) used the VDP to characterize equilibria in silent timing games.

[22] Beyond the applications already cited, Choi and Smith (2017) recently showed that Quah and Strulovici (2012)'s single-crossing aggregation result can be derived from the VDP, and extended it to provide conditions under which a sum of quasiconcave functions is quasiconcave.

Step 1: Necessity. In particular, (10) must hold if $\theta_3 = \theta_2 + \varepsilon$ and $\theta_1 = \theta_2 - \varepsilon$, for any θ_2 and ε. Divide numerator and denominator by ε. In the limit as $\varepsilon \to 0$, this says that $\ell_\theta(s,\theta_2)/\ell_\theta(s,\theta_1)$ increases in s, for any $\theta_1 < \theta_2$; that is, $\ell_\theta(s,\theta)$ must be LSPM.

Step 2: Sufficiency. We must show that ℓ_θ LSPM implies (10). If ℓ is differentiable in θ, then (10) is increasing in s iff its derivative w.r.t. s is positive:

$$(\ell_s(s,\theta_3) - \ell_s(s,\theta_2))\,(\ell(s,\theta_2) - \ell(s,\theta_1)) - (\ell(s,\theta_3) - \ell\,(s,\theta_2))\,(\ell_s(s,\theta_2) - \ell_s(s,\theta_1)) \geq 0$$

Rewrite this as follows, defining $\varepsilon_1 \equiv \theta_2 - \theta_1 > 0$, $\varepsilon_3 \equiv \theta_3 - \theta_2 > 0$, and $\theta_2 \equiv \theta$:

$$(\ell_s(s,\theta+\varepsilon_2) - \ell_s(s,\theta))\,(\ell(s,\theta) - \ell(s,\theta-\varepsilon_1)) \geq (\ell(s,\theta+\varepsilon_2) - \ell(s,\theta))\,(\ell_s(s,\theta) - \ell_s(s,\theta-\varepsilon_1))$$

Now, this holds with equality at $\varepsilon_2 = 0$, so it holds $\forall \varepsilon_2 > 0$ provided the LHS is increasing in ε_2. Taking derivatives w.r.t. ε_2, this yields the sufficient condition:

$$\ell_{s\theta}(s,\theta+\varepsilon_2)\,(\ell(s,\theta) - \ell(s,\theta-\varepsilon_1)) - \ell_\theta(s,\theta+\varepsilon_2)\,(\ell_s(s,\theta) - \ell_s(s,\theta-\varepsilon_1)) \geq 0 \tag{11}$$

But (11) holds with equality at $\varepsilon_1 = 0$, and so to show that it holds $\forall \varepsilon_1$ it is sufficient that the LHS be increasing in ε_1:

$$\ell_{s\theta}(s,\theta+\varepsilon_2)\ell_\theta(s,\theta-\varepsilon_1) \geq \ell_\theta(s,\theta+\varepsilon_2)\ell_{s\theta}(s,\theta-\varepsilon_1)$$

Replacing $\theta + \varepsilon_2$ and $\theta - \varepsilon_1$ with (respectively) θ_2 and $\theta_1 < \theta_2$, this rearranges to $\ell_\theta(s,\theta)$ is LSPM:

$$\frac{\ell_{s\theta}(s,\theta_2)}{\ell_\theta(s,\theta_2)} > \frac{\ell_{s\theta}(s,\theta_1)}{\ell_\theta(s,\theta_1)} \iff \frac{\ell_\theta(s,\theta_2)}{\ell_\theta(s,\theta_1)} \uparrow \text{ in } s \text{ for } \theta_2 > \theta_1$$

A.3 PLRP* in Terms of ϕ: Proof of Theorem 1 (c)

We need to show that for any θ and $s_0 < s$, $\ell_\theta(s,\theta) \equiv \phi(s,\theta)/\phi(s_0,\theta)$ is LSPM iff both ϕ and $\phi^2(\log\phi)_{s\theta}$ are LSPM. Differentiate $\ell(s,\theta)$ in θ to obtain

$$\ell_\theta(s,\theta) = \frac{\phi(s,\theta)}{\phi(s_0,\theta)}\left((\log\phi(s,\theta))_\theta - (\log\phi(s_0,\theta))_\theta\right) \tag{12}$$

Assume ϕ is LSPM, as it is well-known (and clear from (12)) that ϕ is LSPM iff $\ell_\theta \geq 0$.

Step 1. Sufficiency. Assume $\phi^2(\log\phi)_{s\theta}$ is LSPM. To prove that ℓ_θ is LSPM, i.e. that $\ell_\theta(s_2,\theta_2)\ell_\theta(s_1,\theta_1) \geq \ell_\theta(s_1,\theta_2)\ell_\theta(s_2,\theta_1)$ for all $s_2 > s_1$ and $\theta_2 > \theta_1$, it suffices to prove that $\ell_\theta > 0$ implies $(\log \ell_\theta)_{s\theta} \geq 0$. For if $\ell_\theta(s_1, \theta_2) = 0$, the inequality holds trivially by $\ell_\theta \geq 0$. If $\ell_\theta(s_1, \theta_2) > 0$, then also $\ell_\theta(s_2, \theta_2) > 0$ by (12) and LSPM of ϕ. Then $\ell_\theta(s_1, \theta_2)/\ell_\theta(s_2, \theta_2)$ rises (weakly) as $\theta_2 \downarrow \theta_1$ by the MLRP — equivalent to $(\log \ell_\theta)_{\theta s} > 0$ when $\ell_\theta > 0$ — yielding the desired inequality.

So choose $s_0 < s$ and θ with $\ell_\theta(s, \theta) > 0$. Differentiating (12) in s yields

$$(\log \ell_\theta(s,\theta))_s = (\log \phi(s,\theta))_s + \frac{(\log \phi(s,\theta))_{\theta s}}{(\log \phi(s,\theta))_\theta - (\log \phi(s_0,\theta))_\theta}$$

Differentiating in θ yields $(\log \ell_\theta)_{s\theta} = (\log \phi(s,\theta))_{s\theta} / \left((\log \phi(s,\theta))_\theta - (\log \phi(s_0,\theta))_\theta\right)^2 \times$

$$\begin{array}{c} \left((\log \phi(s,\theta))_\theta - (\log \phi(s_0,\theta))_\theta\right)^2 + \left((\log \phi(s,\theta))_\theta - (\log \phi(s_0,\theta))_\theta\right) \frac{(\log \phi(s,\theta))_{\theta\theta s}}{(\log \phi(s,\theta))_{s\theta}} \\ - \left((\log \phi(s,\theta))_{\theta\theta} - (\log \phi(s_0,\theta))_{\theta\theta}\right) \end{array} \tag{13}$$

To show that this is non-negative, the derivative in s_0 is $(\log \phi(s_0,\theta))_{\theta s} \times$

$$\begin{aligned} & -2(\log \phi(s,\theta))_\theta + 2(\log \phi(s_0,\theta))_\theta - \frac{(\log \phi(s,\theta))_{\theta\theta s}}{(\log \phi(s,\theta))_{s\theta}} + \frac{(\log \phi(s_0,\theta))_{\theta\theta s}}{(\log \phi(s_0,\theta))_{\theta s}} \\ = \; & \left(\log \left[(\phi(s_0,\theta))^2 (\log \phi(s_0,\theta))_{s\theta}\right]\right)_\theta - \left(\log \left[(\phi(s,\theta))^2 (\log \phi(s,\theta))_{s\theta}\right]\right)_\theta \end{aligned} \tag{14}$$

This is nonpositive by LSPM of $\phi^2(\log \phi)_{s\theta}$ and $s_0 < s$, implying that expression (13) weakly rises as s_0 falls. Since it vanishes *at* $s_0 = s$, it is non-negative for all $s_0 < s$.

Step 2. Necessity. If $\phi^2(\log \phi)_{s\theta}$ is not LSPM, then we can choose a point (s, θ) where $\left(\log \left(\phi^2(\log \phi)_{s\theta}\right)\right)_\theta$ strictly falls in s. Thus (14) is positive for all $s_0 < s$ in some neighborhood of s. So (13) vanishes at $s_0 = s$ and falls (strictly) as s_0 falls in a neighborhood of s, implying that $\exists s_0 < s$ and θ with $\ell_\theta(s,\theta) > 0$ and (13) negative. So ℓ_θ is not LSPM. □

A.4 Weak VDP: Singular Matrices in Proof of Lemma 1

Assume now the matrix Φ in the §5 proof is singular: columns with indices in $\mathcal{I} = \{i^*, i_1, \ldots, i_n\}$ are linearly dependent. As before (8), toward a contradiction, assume $u(\theta_i) \geq 0$ for $i_m < i < i_{m+1}$ with m odd, and $u(\theta_i) \leq 0$ for $i_m < i < i_{m+1}$ with m even.

Step 1. I first identify linearly independent columns $(\phi_{j,i})_{j \in J}$ with specific properties that will be used to derive a contradiction in Step 2. Start with any linearly independent column vectors $(\phi_{j,i})_{j \in \mathcal{J}}$ with indices in $\mathcal{I}' \subseteq \{i_1, \ldots, i_n\}$. For each $i \notin \mathcal{I}'$, from lowest to highest, replace $\mathcal{I}'$ with $\mathcal{I}' \cup \{i\}$ if two conditions hold: First, column vector $(\phi_{j,i})_{j\in\mathcal{J}}$ is linearly independent from columns with indices $i \in \mathcal{I}'$. Second, if $i \in (i_m, i_{m+1})$ for $m = 1, 2, \ldots, n$, and there are fewer than m indices below i in the set $\mathcal{I}'$ so far. This construction ensures that each $i \in (i_m, i_{m+1}) \setminus \mathcal{I}'$ has *at most* m indices below it in $\mathcal{I}'$.[23] If it has fewer than m indices below it in $\mathcal{I}'$, then i is the index of a column linearly dependent on those in $\mathcal{I}'$, for otherwise it would have been added to $\mathcal{I}'$. Finally, enumerate (say M) indices in $\mathcal{I}'$ as $k_1 < k_2 < \cdots < k_M$, where $M \leq n$.[24]

Step 2. Select $(M+1)$ row indices $\mathcal{J}' \subseteq \mathcal{J}$ such that the matrix $[\phi_{j,i}]$ with $i \in \mathcal{I}'$ and $j \in \mathcal{J}'$ contains at least one nonvanishing $M \times M$ minor (this is possible by linear independence of the M vectors $(\phi_{j,i})_{j\in\mathcal{J}}$ with $i \in \mathcal{I}'$). Assuming wlog that $\mathcal{J}' = \{1, 2, \ldots, M+1\}$, write (7) as follows for $j \in \mathcal{J}'$:

$$\begin{pmatrix} -\left(\alpha_1 - \sum_{i\notin\mathcal{I}'} \phi_{1,i} u(\theta_i)\right) & \phi_{1,k_1} & \cdots & \phi_{1,k_M} \\ \vdots & \vdots & & \vdots \\ -\left(\alpha_{M+1} - \sum_{i\notin\mathcal{I}'} \phi_{M+1,i} u(\theta_i)\right) & \phi_{M+1,k_1} & \cdots & \phi_{M+1,k_M} \end{pmatrix} \begin{pmatrix} 1 \\ u(\theta_{k_1}) \\ \vdots \\ u\left(\theta_{k_M}\right) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

As $M+1$ equations in M unknowns, the determinant of the LHS matrix vanishes:

$$\begin{vmatrix} -\alpha_1 & \phi_{1,k_1} & \cdots & \phi_{1,k_M} \\ \vdots & \vdots & & \vdots \\ -\alpha_{M+1} & \phi_{M+1,k_1} & \cdots & \phi_{M+1,k_M} \end{vmatrix} + \sum_{i\notin\mathcal{I}'} \begin{vmatrix} \phi_{1,i} & \phi_{1,k_1} & \cdots & \phi_{1,k_M} \\ \vdots & \vdots & & \vdots \\ \phi_{M+1,i} & \phi_{M+1,k_1} & \cdots & \phi_{M+1,k_M} \end{vmatrix} u(\theta_i) = 0$$

Now, the first determinant is negative — by $(-1)^{j+1}\alpha_j > 0$ (the weak VDP premise in (7)), total positivity, and linear independence.[25] So the sum must be positive. But for $i \notin \mathcal{I}'$, the ith summand is nonpositive: By Step 1, either (a) column i is a linear combination of the remaining columns, so that the matrix is singular, or (b) $i \in (i_m, i_{m+1})$ has exactly m indices below it in $\mathcal{I}'$, and thus the determinant has sign $(-1)^m$ — weakly opposite to the assumed sign of $u(\theta_i)$. A contradiction. □

[23] Since $\mathcal{I}'$ starts out as a subset of $\{i_1, \ldots, i_n\}$ (with exactly m indices below i), and only adds indices $i \in (i_m, i_{m+1})$ with fewer than m terms below them in $\mathcal{I}'$.

[24] In particular, the highest index $i > i_n$ has at most n terms $\{k_1, \ldots, k_M\}$ below it in $\mathcal{I}'$.

[25] Expanding along the first column, the determinant is $\sum_{j=1}^{M+1}(-1)^{j+2}\alpha_j A_{j,1}$, where $A_{j,1}$ is the minor obtained by deleting row j and column 1; each minor is nonnegative by total positivity, and at least one is nonvanishing by the linearly independent selection of rows and columns.

A.5 Location Families: Proof of Corollary 1

Define $\phi(s,\theta) = \psi(s-\theta)$. Then $(\log\phi(s,\theta))_s \equiv (\log\psi(s-\theta))_s \equiv \delta(s-\theta)$, so that $(\log\phi)_{s\theta} = -\delta'(s-\theta)$. Then ϕ obeys the MLRP — i.e. is LSPM — iff $\delta' \leq 0$. By Theorem 1(c), ϕ obeys the PLRP iff additionally $\phi^2 \cdot (\log\phi)_{s\theta}$ is itself log-supermodular, requiring

$$\frac{\partial^2\left(2\log\phi + \log\left((\log\phi)_{s\theta}\right)\right)}{\partial s\partial\theta} \geq 0$$

and therefore

$$\frac{\partial}{\partial\theta}2(\log\phi)_s + \frac{\partial^2}{\partial s\partial\theta}\log(-\delta'(s-\theta)) \equiv \frac{\partial}{\partial\theta}2\delta(s-\theta) + \frac{\partial}{\partial\theta}\frac{\delta''(s-\theta)}{\delta'(s-\theta)} \geq 0$$

i.e. $2\delta + \delta''/\delta'$ must be weakly decreasing.

A.6 Truncations: Proof of Theorem 2

Let ϕ be a TP-3 function with $\phi(\theta,\theta) = 0$. We want to show that for any $\theta_1 < \theta_2 < \theta_3$ and $s_1 < s_2 < s_3$:

$$\begin{vmatrix} \phi_{11}\mathbb{I}_{\theta_1 \leq s_1} & \phi_{12}\mathbb{I}_{\theta_2 \leq s_1} & \phi_{13}\mathbb{I}_{\theta_3 \leq s_1} \\ \phi_{21}\mathbb{I}_{\theta_1 \leq s_2} & \phi_{22}\mathbb{I}_{\theta_2 \leq s_2} & \phi_{23}\mathbb{I}_{\theta_3 \leq s_2} \\ \phi_{31}\mathbb{I}_{\theta_1 \leq s_3} & \phi_{32}\mathbb{I}_{\theta_2 \leq s_3} & \phi_{33}\mathbb{I}_{\theta_3 \leq s_3} \end{vmatrix} \geq 0 \tag{15}$$

where $\phi_{ji} \equiv \phi(s_j,\theta_i)$. If all indicators are 1, then ϕ TP-3 gives this. If two or more indicators vanish, the first indicator inequality to fail is $\theta_3 \leq s_1$, followed by $\theta_3 \leq s_2$ or $\theta_2 \leq s_1$. So either the last two entries in the first row vanish, or the first two entries in the last column vanish. Either way, the determinant is non-negative by ϕ TP-2.

So we only need to verify TP-3 in the case where exactly one indicator equals zero, in which case it will be the top right entry $\mathbb{I}_{\theta_3 \leq s_1}$. Then (15) reduces to:

$$0 \;\leq\; \begin{vmatrix} \phi_{11} & \phi_{12} & 0 \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{vmatrix} = \phi_{11}\phi_{22}\phi_{33} - \phi_{11}\phi_{23}\phi_{32} - \phi_{12}\phi_{21}\phi_{33} + \phi_{12}\phi_{31}\phi_{23} \tag{16}$$

So $(\phi_{11}\phi_{22} - \phi_{12}\phi_{21})\,\phi_{33} \geq \phi_{23}\,(\phi_{11}\phi_{32} - \phi_{12}\phi_{31})$. Both sides of this inequality are non-negative by the fact that ϕ is LSPM, so we can rearrange it as follows:

$$\frac{\phi_{22}\phi_{33} - \frac{\phi_{12}}{\phi_{11}}\phi_{21}\phi_{33}}{\phi_{32}\phi_{23} - \frac{\phi_{12}}{\phi_{11}}\phi_{31}\phi_{23}} \geq 1 \tag{17}$$

The LHS of (17) is *decreasing in* $\frac{\phi_{12}}{\phi_{11}}$ whenever the following is nonpositive:

$$\phi_{31}\phi_{23}\left(\phi_{22}\phi_{33}-\frac{\phi_{12}}{\phi_{11}}\phi_{21}\phi_{33}\right)-\phi_{21}\phi_{33}\left(\phi_{32}\phi_{23}-\frac{\phi_{12}}{\phi_{11}}\phi_{31}\phi_{23}\right)=\phi_{23}\phi_{33}\left(\phi_{31}\phi_{22}-\phi_{32}\phi_{21}\right)$$

which holds by LSPM: $\phi_{32}/\phi_{22} \geq \phi_{31}/\phi_{21}$. So (17), and hence (16), are tightest at the *maximum* value of $\phi_{12}/\phi_{11} \equiv \phi(s_1, \theta_2)/\phi(s_1, \theta_1)$. By LSPM, this ratio is increasing in s_1, and so its upper bound is attained at $s_1 = \theta_3$, by the assumed failure of the inequality $\theta_3 \leq s_1$. So to verify TP-3, we need only check (16) *at* $s_1 = \theta_3$. But here, by our assumption $\phi(\theta, \theta) \equiv 0$, we have $\phi_{13} \equiv \phi(s_1, \theta_3) = 0$. So the indicator $\mathbb{I}_{\theta_3 \leq s_1}$ does not affect the determinant (16). Then the determinant is non-negative, since ϕ is TP-3. □

A.7 Stochastic Shifts: Proof of Corollary 3(*b*)

Assume ϕ obeys the PLRP. By Corollary 2(*b*), the following is quasiconcave for any quasiconcave $u : \Theta \to \mathbb{R}$:

$$B_u(s) = \int \phi(s, \theta) u(\theta) d\mu(\theta)$$

To prove that $-\Phi_s(s, \theta)$ is LSPM, I construct a quasiconcave function $u(\cdot)$ for which quasiconcavity of B_u would otherwise fail.

To this end, define $\psi(s, \theta) \equiv -\Phi_s(s, \theta)$. If ψ is not LSPM, then $(\log \psi(s, \theta))_{\theta s} < 0$ at some (s, θ), i.e. $\psi_\theta(s, \theta)/\psi(s, \theta)$ is strictly decreasing in s at (s, θ). By continuity, this holds also near (s, θ), and so we can choose $s_1 < s_2 < s_3$ and $\theta_1 < \theta_2$ such that

$$\frac{\psi_\theta(s_1,\theta)}{\psi(s_1,\theta)} > \frac{\psi_\theta(s_2,\theta)}{\psi(s_2,\theta)} > \frac{\psi_\theta(s_3,\theta)}{\psi(s_3,\theta)} \qquad \forall \theta \in [\theta_1, \theta_2]$$

So $\psi(s_2, \theta)/\psi(s_1, \theta)$ is strictly decreasing in θ on $[\theta_1, \theta_2]$, by the first inequality, as is $\psi(s_3, \theta)/\psi(s_2, \theta)$ by the second inequality. Let $u : \Theta \to \mathbb{R}$ be constant outside $[\theta_1, \theta_2]$, $u'(\theta) = 1$ in the left half of the interval, and in the right half of the interval

$$u'(\theta) = -\left(\int_{\theta_1}^{\frac{\theta_1+\theta_2}{2}} \psi(s_2, t) dt\right) \Big/ \left(\int_{\frac{\theta_1+\theta_2}{2}}^{\theta_2} \psi(s_2, t) dt\right)$$

As u' is downcrossing, u is quasiconcave. Integrating by parts, $B'_u(s_2) = \int \psi(s_2, \theta) u'(\theta) d\theta$, which vanishes by construction. Since $\psi(s_1, \theta)/\psi(s_2, \theta)$ is increasing on $[\theta_1, \theta_2]$, we have $B'_u(s_1) = \int_{\theta_1}^{\theta_2} \frac{\psi(s_1,\theta)}{\psi(s_2,\theta)} \psi(s_2, \theta) u'(\theta) d\theta < 0$.[26] Since $\psi(s_3, \theta)/\psi(s_2, \theta)$ is increasing, the same

[26] By Karlin and Rubin (1956), if $b(x)$ upcrosses at x_0 and $a(x)$ is nondecreasing, then $\int_{-\infty}^{x_0} a(x)b(x)dx + \int_{x_0}^{\infty} a(x)b(x)dx \geq a(x_0) \int b(x)dx$, strictly so if $a(x)$ is increasing on an interval with

logic yields $B'_u(s_3) > 0$. Then $B'_u(s_1) < 0 < B'_u(s_3)$ for $s_1 < s_2 < s_3$, and so B'_u has a strictly *upcrossing* portion, a contradiction.

For the strict claim, assume for a contradiction that ϕ obeys the strict PLRP, but $\psi(s,\theta) \equiv -\Phi_s(s,\theta)$ obeys only the weak MLRP. So $(\log\psi(s,\theta))_{\theta s} = 0$ at some (θ, s), and so again by continuity, we can choose $s_1 < s_2$ and $\theta_1 < \theta_2$ so that for all $s \in [s_1, s_2]$, $\psi(s,\theta)/\psi(s_2,\theta)$ is constant in θ on $[\theta_1, \theta_2]$. Now choose the same function $u(\cdot)$ as above; again by construction, $B'_u(s_2) = 0$, and so for all $s \in [s_1, s_2]$, $\psi(s,\theta)/\psi(s_2,\theta)$ constant implies that $B_u(s)$ is constant on $[s_1, s_2]$, contradicting Corollary 2(c). □

A.8 Amplification: Proof of Theorem 3

Given a stochastic kernel (ϕ, μ) and functions $u, v : \Theta \to \mathbb{R}$ with $v' \geq 0$, integrate $U(s) = \int \phi(s,\theta)u(\theta)d\mu(\theta)$ by parts to obtain $U'(s) = \int -\Phi_s(s,\theta)u'(\theta)d\theta$. Similarly, if $V(s) = \int \phi(s,\theta)u(\theta)d\mu(\theta)$, then $V'(s) = \int -\Phi_s(s,\theta)v'(\theta)d\theta$. Define $\psi(\theta, s) \equiv -\Phi_s(\theta, s)$, and assume that ψ is LSPM (by Corollary 3(b), this holds if ϕ obeys the PLRP). Now choose $s_2 > s_1$ and write $U'(s_2)V'(s_1) - U'(s_1)V'(s_2)$ as follows:[27]

$$
\begin{aligned}
&\left[\int \psi(\theta,s_2)u'(\theta)d\theta\right]\left[\int \psi(\theta,s_1)v'(\theta)d\theta\right] - \left[\int \psi(\theta,s_1)u'(\theta)d\theta\right]\left[\int \psi(\theta,s_2)v'(\theta)d\theta\right] \\
= &\int\int \frac{\psi(r,s_2)}{\psi(r,s_1)}\left[u'(r)v'(\theta) - v'(r)u'(\theta)\right]\psi(r,s_1)\psi(\theta,s_1)drd\theta \\
= &\int\int \frac{\psi(\theta,s_2)}{\psi(\theta,s_1)}\left[u'(\theta)v'(r) - v'(\theta)u'(r)\right]\psi(\theta,s_1)\psi(r,s_1)drd\theta \\
= &\frac{1}{2}\int\int \left(\frac{\psi(r,s_2)}{\psi(r,s_1)} - \frac{\psi(\theta,s_2)}{\psi(\theta,s_1)}\right)\left(\frac{u'(r)}{v'(r)} - \frac{u'(\theta)}{v'(\theta)}\right)v'(r)v'(\theta)\psi(r,s_1)\psi(\theta,s_1)drd\theta
\end{aligned}
$$

By LSPM, $\left(\frac{\psi(r,s_2)}{\psi(r,s_1)} - \frac{\psi(\theta,s_2)}{\psi(\theta,s_1)}\right)$ is nonnegative if $r > \theta$ and nonpositive if $r < \theta$. If u is a convex transformation of v, then $\left(\frac{u'(r)}{v'(r)} - \frac{u'(\theta)}{v'(\theta)}\right) \gtreqless 0$ iff $r \gtrless \theta$, and so $U'(s_2)V'(s_1) \geq U'(s_1)V'(s_2)$ for all $s_2 > s_1$. So U is more convex than V. If u is a concave function of v, the integrand is everywhere nonpositive, and so U is more concave than V. □

$b(x) \neq 0$. Then $\int a(x)b(x)dx > 0$ if strict inequality holds and $\int b(x)dx = 0$, or if $\int b(x)dx > 0$ and $a(x_0) > 0$. Ditto, $\int a(x)b(x)dx \leq 0$ if $a(x) \geq 0$ is increasing while $b(x)$ is downcrossing with $\int b(x)dx \leq 0$.

[27]This algebraic argument mimics the proof of Tchebyshev's inequality. An alternative geometric proof, following Karlin (1968), uses the fact that a function is concave (convex) iff it crosses any line at most twice, and if exactly twice, then initially from below (above, respectively). So if $u : \Theta \to \mathbb{R}$ is a concave transformation of (increasing) $v : \Theta \to \mathbb{R}$, then for any constants $\alpha, \beta \in \mathbb{R}$ the function $u(\cdot) - (\alpha u(\cdot) + \beta)$ is at most 2-crossing, and if exactly 2-crossing, then initially from below. So by the VDP, $U(s) - (\alpha V(s) + \beta) = \int \phi(s,\theta)(u(\theta) - (\alpha v(\theta) + \beta))d\theta$ is also at most 2-crossing, and if exactly 2-crossing, then initially from below. That is, U is a concave transformation of V.

References

Chade, H., and J. Swinkels (2019): "The No-Upward-Crossing Condition, Comparative Statics, and the Moral-Hazard Problem," ASU and NWU mimeo.

Choi, M., and L. Smith (2017): "Ordinal Aggregation Results via Karlin's Variation Diminishing Property," *Journal of Economic Theory*, 168, 1–11.

Diamond, P., and J. Stiglitz (1974): "Increases in Risk and in Risk Aversion," *Journal of Economic Theory*, 8(3), 337–360.

Holmstrom (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10(1), 74–91.

Jewitt, I. (1987): "Risk Aversion and the Choice Between Risky Prospects: The Preservation of Comparative Statics Results," *Review of Economic Studies*, pp. 73–85.

——— (1988): "Justifying the First-Order Approach to Principal-Agent Problem," *Econometrica*, 56(5), 1177–1190.

Karlin, S. (1957): "Polya Type Distributions, II," *The Annals of Mathematical Statistics*, 28(2), 281–308.

——— (1968): *Total Positivity: Volume 1*. Stanford University Press, Palo Alto, CA.

Karlin, S., and H. Rubin (1956): "The Theory of Decision Problems for Distributions with the Monotone Likelihood Ratio," *Annals of Mathematical Statistics*, 27(2), 272–299.

Milgrom, P. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), 380–391.

Milgrom, P., and C. Shannon (1994): "Monotone Comparative Statics," *Econometrica*, 62(1), 157–180.

Park, A., and L. Smith (2008): "Caller Number Five and Related Timing Games," *Theretical Economics*, (3), 231–256.

Pratt, J. (1988): "Aversion to One Risk In the Presence of Others," *Journal of Risk and Uncertainty*, 1, 395–413.

Quah, J., and B. Strulovici (2012): "Aggregating the Single Crossing Property," *Econometrica*, 80(5), 2333–2348.