Online Appendix

to

Information Validates the Prior: A Theorem on Bayesian Updating and Applications

Navin Kartik* Frances Xu Lee† Wing Suen‡

July 21, 2020§

*Department of Economics, Columbia University. E-mail: nkartik@columbia.edu.

†Quinlan School of Business, Loyola University Chicago. E-mail: francesxu312@gmail.com.

‡Faculty of Business and Economics, The University of Hong Kong. E-mail: wsuen@econ.hku.hk.

§Cosmetic changes on January 15, 2021 after journal's page proofs.

B. Supplementary Appendix

B.1. Discussion of Theorem 1

The following example shows that even with likelihood-ratio ordered priors, the "direction" portion of Theorem 1 can fail with a non-MLRP experiment.

Example B.1. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$, where $\omega_1 < \omega_2 < \omega_3$. Consider the following non-MLRP experiment $\mathcal{E}$ with signal space $S = \{s_1, s_2\}$:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

where the entry in row r and column c is $\Pr(s_r|\omega_c)$. Consider priors $\beta_A = (1, 0, 0) <_{LR} \beta_B = (0, x, 1-x)$ for any $x \in (0, 1)$. Plainly, $m_B^{s_2} = \omega_3$, and hence $\omega_1 = m_A < m_B < \omega_3 = \mathbb{E}_A^{\mathcal{E}}[m_B^s]$. $\diamond$

The priors in Example B.1 violate full support, but the point goes through if β_A and β_B are perturbed to satisfy full support. Alonso and Câmara (2016, pp. 674–675) use a similar example to illustrate how a "skeptic" can design information to persuade a "believer."

B.2. The role of linearity and MLRP in the signaling application

Consider the costly signaling application from Section III. Recall that in the LCSE, the sender's strategy $\rho(\cdot)$ is determined by the initial condition $\rho(0) = 0$ and the differential equation (5):

$$\frac{\partial c(\rho(t), t)}{\partial r}\rho'(t) = \frac{\partial \mathbb{E}_{s|t}\left[\beta(s;t)\right]}{\partial \pi}.$$

Lemma 3 established that

$$\frac{\partial \mathbb{E}_{\tilde{s}|t}\left[\beta(\tilde{s};t)\right]}{\partial \pi} \leq \frac{\partial \mathbb{E}_{s|t}\left[\beta(s;t)\right]}{\partial \pi} \tag{B.1}$$

when signal $\tilde{s}$ is more informative (i.e., drawn from a more informative experiment) than signal s. Inequality (B.1) implies that the solution to the aforementioned initial-value problem is pointwise lower under the more informative experiment, and hence the equilibrium signaling level $\rho(t)$ is lower for every type when the receiver has access to $\tilde{s}$ rather than s.

We show below how the conclusion can be altered by dropping either linearity of the sender's payoff in the receiver's posterior (Example B.2) or the MLRP of the receiver's experiments (Example B.3).

Example B.2. Letting $V(\beta) \equiv \beta/(1-\beta)$, suppose the sender's payoff is

$$V(\beta) - c(r,t),$$

which is convex in the receiver's posterior β. Condition (4) in Section III continues to imply the relevant single-crossing condition for this modified objective. Using Bayes rule, we compute

$$\mathbb{E}_{s|t}[V(\beta(s;\pi))] = \frac{\pi}{1-\pi}\mathbb{E}_{s|t}\left[\frac{g(s|1)}{g(s|0)}\right].$$

Differentiating and evaluating at $\pi = t$,

$$\frac{\partial \mathbb{E}_{s|t}[V(\beta(s;t))]}{\partial \pi} = \frac{1}{t(1-t)}\mathbb{E}_{s|t}\left[\frac{\beta(s;t)}{1-\beta(s;t)}\right].$$

The term inside the expectation operator on the right-hand side above is a convex function of $\beta(\cdot)$. It follows that

$$\frac{\partial \mathbb{E}_{\tilde{s}|t}[V(\beta(\tilde{s};t))]}{\partial \pi} \geq \frac{\partial \mathbb{E}_{s|t}[V(\beta(s;t))]}{\partial \pi},$$

by contrast to (B.1). That is, the convexity in $V(\cdot)$ is strong enough to ensure that the marginal benefit from inducing a higher interim belief π (locally, at $\pi = t$) is higher when the exogenous signal is more informative. It follows that in the LCSE, all types bear a *higher* signaling cost when the exogenous signal is more informative.[1] ◇

Example B.3. To see that MLRP-experiments are important, we have to modify the signaling model of Section III by introducing more states, because any experiment in a two-state model satisfies MLRP.

Assume a full-support common prior about the state $\omega \in \{0,1,2\}$. The sender receives some private information, indexed by $t \in [0,1]$, which updates his belief about the state to $(z, 1-z(1+t), zt)$, where each element of this vector is the probability assigned to the corresponding state. The parameter $z \in (0,1/2)$ is a commonly-known constant. We refer to t as the sender's type. Letting $M(\beta) \equiv \sum_\omega \omega\beta(\omega)$ be the receiver's expectation of the state when

[1] On the other hand, if $V(\beta) \equiv \log[\beta/(1-\beta)]$, then the local marginal benefit of inducing a higher interim belief is independent of the exogenous experiment. The reason is that $V(\beta(s,\pi)) = \log\left(\frac{\pi}{1-\pi}\right) + \log\left(\frac{g(s|1)}{g(s|0)}\right)$ and hence $\partial\mathbb{E}_{s|t}[V(\beta(s;t))]/\partial\pi$ does not depend on $g(\cdot)$. Note that $V(\cdot)$ here is neither convex nor concave.

she holds belief β, the sender's payoff is

$$M(\beta) - c(r, t).$$

Let s represent the outcome of an uninformative experiment, and let $\beta^s_{\hat{t}}$ represent the posterior of the receiver after observing s when she puts probability one on the sender's type $\hat{t}$. It clearly holds that

$$\mathbb{E}_{s|t}\left[M(\beta^s_{\hat{t}})\right] = M(\beta^s_{\hat{t}}) = 1 - z + z\hat{t}.$$

The derivative with respect to $\hat{t}$, evaluated at $\hat{t} = t$, is

$$\frac{\partial \mathbb{E}_{s|t}\left[M(\beta^s_t)\right]}{\partial \hat{t}} = z. \tag{B.2}$$

Now consider an informative experiment with a binary signal space, $\tilde{s} \in \{l, h\}$. Let the probability distributions $g(\tilde{s}|\omega)$ be given by:

$$\begin{bmatrix} g(l|0) & g(l|1) & g(l|2) \\ g(h|0) & g(h|1) & g(h|2) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

This experiment is the same as that in Example B.1 of Supplementary Appendix B.1; it does not have the MLRP.

Suppose the receiver ascribes probability one to the sender's type $\hat{t}$. By Bayes rule, if the signal realization is $\tilde{s} = l$, the receiver's posterior is $\beta^l_{\hat{t}} = (0, 1, 0)$, with $M(\beta^l_{\hat{t}}) = 1$. For signal realization $\tilde{s} = h$,

$$\beta^h_{\hat{t}} = \left(\frac{1}{1+\hat{t}}, 0, \frac{\hat{t}}{1+\hat{t}}\right), \quad \text{with} \quad M(\beta^h_{\hat{t}}) = \frac{2\hat{t}}{1+\hat{t}}.$$

The sender of type t's expectation is

$$\mathbb{E}_{\tilde{s}|t}\left[M(\beta^{\tilde{s}}_{\hat{t}})\right] = (1 - z(1+t))M(\beta^l_{\hat{t}}) + z(1+t)M(\beta^h_{\hat{t}}).$$

The derivative with respect to $\hat{t}$, evaluated at $\hat{t} = t$, is

$$\frac{\partial \mathbb{E}_{\tilde{s}|t}\left[M(\beta^{\tilde{s}}_t)\right]}{\partial \hat{t}} = \frac{2z}{1+t}. \tag{B.3}$$

Combining (B.2) and (B.3),

$$\frac{\partial \mathbb{E}_{\tilde{s}|t}\left[M(\beta^{\tilde{s}}_t)\right]}{\partial \hat{t}} \geq \frac{\partial \mathbb{E}_{s|t}\left[M(\beta^s_t)\right]}{\partial \hat{t}}, \tag{B.4}$$

which is the opposite inequality to (B.1), even though $\tilde{s}$ is drawn a more informative experiment than s.

In this example, both $\mathbb{E}_{s|t}[M(\beta(s;\hat{t}))]$ and $\mathbb{E}_{\tilde{s}|t}[M(\beta(\tilde{s};\hat{t}))]$ are supermodular in the sender's type t. The assumption that $\partial c(r,t)/\partial r\partial t < 0$ ensures that indifference curves in the space of $(r,\hat{t})$ for different types are single crossing. As local incentive compatibility then implies global incentive compatibility, (B.4) implies that in the LCSE all types incur higher signaling costs when the receiver has access to the more informative experiment. $\diamond$

References

Alonso, Ricardo, and Odilon Câmara. 2016. "Bayesian Persuasion with Heterogeneous Priors." *Journal of Economic Theory* 165 (1): 672–706.