Electronic Companion - "Human and Machine: The Impact of Machine Input on Decision-Making Under Cognitive Limitations"

Tamer Boyacı
ESMT Berlin, 10178 Berlin, Germany, tamer.boyaci@esmt.org

Caner Canyakmaz
Faculty of Business, Ozyegin University, Istanbul 34794, Turkey, caner.canyakmaz@ozyegin.edu.tr

Francis de Véricourt
ESMT Berlin, 10178 Berlin, Germany, francis.devericourt@esmt.org

---

Appendix A: Proofs of Results

This section contains proofs of all results in the main paper and the appendices.

Proof of Lemma 1 (4) follows from Theorem 1 in Matějka and McKay (2015) which are obtained for action $a = y$. $A^\star$ and $C^\star$ are by definition.

Proof of Theorem 1 By (4), we have $p = (1-\mu)\frac{p}{p+(1-p)e^{1/\lambda}} + \mu\frac{pe^{1/\lambda}}{pe^{1/\lambda}+1-p}$ which gives

$$\overline{p} = \frac{\mu}{1-e^{-1/\lambda}} - \frac{1-\mu}{e^{1/\lambda}-1}.$$

Choice probability is 1 when $\overline{p} \geq 1$, or equivalently, $\mu \geq \overline{\mu} = \frac{e^{1/\lambda}}{e^{1/\lambda}+1}$. Similarly, choice probability is 0 when $\overline{p} \leq 0$ or equivalently $\mu \leq \underline{\mu} = \frac{1}{e^{1/\lambda}+1}$. As $\overline{p}$ is linearly increasing in μ with first order derivative $\frac{1}{1-e^{-1/\lambda}} + \frac{1}{e^{1/\lambda}-1} > 0$, $p^\star$ is increasing in μ. Finally, $\overline{\mu}$ is decreasing in λ as $\frac{d\overline{\mu}}{d\lambda} = -\frac{1}{\lambda^2} e^{\frac{1}{\lambda}} \frac{1}{\left(e^{\frac{1}{\lambda}}+1\right)^2} < 0$ and $\underline{\mu}$ is increasing in λ as $\frac{d\underline{\mu}}{d\lambda} = \frac{1}{\lambda^2}\frac{e^{\frac{1}{\lambda}}}{\left(e^{\frac{1}{\lambda}}+1\right)^2} > 0$.

Proof of Corollary 1 $\mu \leq \underline{\mu} = \frac{1}{e^{1/\lambda}+1} \Leftrightarrow \left(\log\frac{1-\mu}{\mu}\right)^{-1} \leq \lambda$ which yields an $a = n$ decision by Theorem 1. Note that $\log\frac{1-\mu}{\mu}$ is positive when $\mu < 0.5$. Similarly, $\mu \geq \overline{\mu} = \frac{e^{1/\lambda}}{e^{1/\lambda}+1} \Leftrightarrow \lambda \geq \left(\log\frac{\mu}{1-\mu}\right)^{-1}$ which leads to $a = y$ decision and the log term is positive when $\mu > 0.5$. Therefore, for any $\mu \neq 0.5$, $\overline{\lambda}$ can be written in absolute terms, that is, $\overline{\lambda} = \left|\log\frac{1-\mu}{\mu}\right|^{-1}$. Then (8) follows. Furthermore, $\frac{d}{d\lambda}p^\star(\lambda) = \frac{1}{\lambda^2}\frac{e^{-\frac{1}{\lambda}}}{\left(e^{-\frac{1}{\lambda}}-1\right)^2}(2\mu-1)$ which is positive when $\mu > 0.5$ and negative when $\mu < 0.5$. Hence the monotonicity result follows.

Proof of Corollary 2 We can write the DM's accuracy in (5) in terms of optimal posterior beliefs that she constructs as

$$A^\star = (1-\mu)(1-p_b) + \mu p_g = \gamma(b|n)(1-p^\star) + \gamma(g|y)p^\star$$

where $\gamma(\omega|a)$ denotes the optimal posterior that the state is ω given action a. When $\mu < \underline{\mu}$, $A^\star = 1-\mu$ as $p^\star = 0$ and $\gamma(b|n) = 1-\mu$. Similarly, when $\mu > \overline{\mu}$, $A^\star = \mu$ as $p^\star = 1$ and $\gamma(g|y) = \mu$. For the case where $\mu \in [\underline{\mu}, \overline{\mu}]$, we use the optimal posterior characterizations that are given in Lemma C.1 (in Appendix C) for $\delta = 1$, which yields $\gamma(g|y) = \overline{\mu} = \frac{e^{1/\lambda}}{e^{1/\lambda}+1}$ and $\gamma(g|n) = \underline{\mu} = \frac{1}{e^{1/\lambda}+1}$. Note that $\gamma(g|y) = \gamma(b|n)$ and we have $A^\star = \overline{\mu} = \frac{e^{1/\lambda}}{e^{1/\lambda}+1}$.

Using the symmetry of mutual information (see Cover and Thomas 2012), we can write (5) as

$$C^\star = \lambda \left[H(\mu) - pH(\gamma(g|y)) - (1-p)H(\gamma(g|n))\right].$$

Assume that $\mu \in [\underline{\mu}, \overline{\mu}]$. Then, as $\gamma(g|y) = 1 - \gamma(g|n)$ we have $H(\gamma(g|y)) = H(\gamma(g|n))$ from the symmetry of the entropy function H in $[0,1]$. Then, $C^\star$ becomes

$$\begin{aligned} C^\star &= \lambda \left[H(\mu) - H\left(\frac{e^{1/\lambda}}{e^{1/\lambda}+1}\right)\right] = \lambda \left[H(\mu) + \frac{e^{1/\lambda}}{e^{1/\lambda}+1}\log\frac{e^{1/\lambda}}{e^{1/\lambda}+1} + \frac{1}{e^{1/\lambda}+1}\log\frac{1}{e^{1/\lambda}+1}\right] \\ &= \lambda \left[H(\mu) + \frac{1}{\lambda}\frac{e^{1/\lambda}}{e^{1/\lambda}+1} - \log\left(e^{1/\lambda}+1\right)\right]. \end{aligned}$$

When, $\mu \notin [\underline{\mu}, \overline{\mu}]$, $C^\star = 0$, as $\gamma(g|y) = \gamma(g|n) = \mu$. Finally, $V^\star$ is found by taking the difference $A^\star - C^\star$.

Proof of Corollary 3 When $\mu \le \underline{\mu}$, $p_b = p_g = 0$, hence $\alpha^\star = 0$ and $\beta^\star = \mu$ by (6). Similarly, when $\mu > \overline{\mu}$, $p_b = p_g = 1$, hence $\alpha^\star = 1-\mu$ and $\beta^\star = 0$. Now assume $\mu \in [\underline{\mu}, \overline{\mu}]$. Writing (6) in terms of optimal posteriors in Lemma C.1 for $\delta = 1$ and plugging in optimal choice in (7), we obtain

$$\begin{aligned} \alpha^\star &= (1-\mu)p_b = (1-\gamma(g|y))p = \frac{1}{e^{1/\lambda}+1}\left(\frac{\mu}{1-e^{-1/\lambda}} - \frac{1-\mu}{e^{1/\lambda}-1}\right) = \frac{\mu\left(e^{1/\lambda}+1\right)-1}{e^{2/\lambda}-1} \\ \beta^\star &= \mu(1-p_g) = \gamma(g|n)(1-p) = \frac{1}{e^{1/\lambda}+1}\left(1 - \frac{\mu}{1-e^{-1/\lambda}} + \frac{1-\mu}{e^{1/\lambda}-1}\right) = \frac{e^{1/\lambda} - \mu\left(e^{1/\lambda}+1\right)}{e^{2/\lambda}-1}. \end{aligned}$$

Proof of Lemma 2 Using $\mu^- = 0$, the result directly follows by (7), Corollary 2 and Corollary 3 for $p^\star$, $\alpha^\star$, $\beta^\star$, $A^\star$, $C^\star$, $V^\star$.

Proof of Proposition 1 $A^\star$ given in 2 is convex in μ for $[0,1]$. Then, by Jensen's inequality,

$$A^\star(\mu) = A^\star\left(\left(1-\frac{\mu}{\mu}\right)\mu^- + \frac{\mu}{\mu^+}\mu^+\right) \le \left(1-\frac{\mu}{\mu^+}\right)A^\star(\mu^-) + \frac{\mu}{\mu^+}A^\star\left(\mu^+\right) = 1 - \frac{\mu}{\mu^+} + \frac{\mu}{\mu^+}A^\star\left(\mu^+\right) = A^\star_m(\mu).$$

Similarly, the value function $V^\star$ is also convex in μ. To see this, note first that $V(\mu)$ is linearly decreasing in $[0, \underline{\mu}]$, convex in $[\underline{\mu}, \overline{\mu}]$ (since the entropy function H is concave) and linearly increasing in $[\overline{\mu}, 1]$. Furthermore, the slope of $\lambda\left[\log\left(e^{\frac{1}{\lambda}}+1\right) - H(\mu)\right]$ is the same at both of these cutoff points. More specifically,

$$\frac{d}{d\mu}_{|\mu=\underline{\mu}} \lambda\left[\log\left(e^{\frac{1}{\lambda}}+1\right) - H(\mu)\right] = \lambda\log\frac{\mu}{1-\mu}_{|\mu=\underline{\mu}} = \lambda\log\frac{\frac{1}{e^{1/\lambda}+1}}{1-\frac{1}{e^{1/\lambda}+1}} = -1.$$

Similarly, $\lambda\log\frac{\overline{\mu}}{1-\overline{\mu}} = 1$. Since the slope is increasing in μ, $V(\mu)$ is convex in μ for $\mu \in [0,1]$. By Jensen's inequality $V^\star(\mu) \le V^\star_m(\mu)$.

Proof of Theorem 2 Note that *i*), *ii*), *iii*) and *vi*) follow by the optimal choice probability in (7) and the fact that $p_m^\star = \mu/\mu^+ p^\star(\mu^+)$.

iv) Using (7) and $\frac{\mu}{\mu^+} < 1$, we have

$$p_m^\star = \frac{\mu}{\mu^+}\left(\frac{\mu^+}{1-e^{-1/\lambda}} - \frac{1-\mu^+}{e^{1/\lambda}-1}\right) = \frac{\mu}{1-e^{-1/\lambda}} - \frac{\frac{\mu}{\mu^+}-\mu}{e^{1/\lambda}-1} > \frac{\mu}{1-e^{-1/\lambda}} - \frac{1-\mu}{e^{1/\lambda}-1} = p^\star$$

v) $p_m^\star > p^\star$ $\frac{\mu}{\mu^+} > \frac{\mu}{1-e^{-1/\lambda}} - \frac{1-\mu}{e^{1/\lambda}-1}$ which can equivalently be written as

$$\frac{1}{e^{1/\lambda}-1} > \mu^+\left(\frac{1}{1-e^{-1/\lambda}} + \frac{1}{e^{1/\lambda}-1} - \frac{1}{\mu^+}\right). \tag{A.1}$$

The right hand side is always positive since $\mu^+ > \frac{e^{1/\lambda}}{e^{1/\lambda}+1}$. That is,

$$\frac{1}{1-e^{-1/\lambda}} + \frac{1}{e^{1/\lambda}-1} - \frac{1}{\mu^+} < 0 \Leftrightarrow \mu^+ > \frac{e^{1/\lambda}-1}{e^{1/\lambda}+1}$$

which is always true since $\mu^+ > \frac{e^{1/\lambda}}{e^{1/\lambda}+1}$. Then, $(A.1)$ can be written as

$$\mu^+ < \frac{\frac{1}{e^{1/\lambda}-1}}{\frac{1}{1-e^{-1/\lambda}} + \frac{1}{e^{1/\lambda}-1} - \frac{1}{x}} = \left(e^{1/\lambda}+1-\frac{e^{1/\lambda}-1}{\mu^+}\right)^{-1} = \widehat{\mu}_c.$$

Note that $\widehat{\mu}_c$ is decreasing in μ^+ and for $\mu^+ = 1$, $\widehat{\mu}_c = 0.5$. Then $\widehat{\mu}_c \geq 0.5$.

Proof of Corollary 4 Assume $\mu \in [\underline{\mu}, \overline{\mu}]$ for a fixed λ. By Theorem 2, since $\widehat{\mu}_c \geq 0.5$, $p_m^\star \geq p^\star$ for $\mu \leq 0.5$. For $\mu < \underline{\mu}$, $p_m^\star \geq p^\star$ by *i*, *ii* and *iii*. This proves the first part. When $\mu > 0.5$, $p_m^\star \leq p^\star$ if $\mu \geq \widehat{\mu}_c$ for a fixed λ. Using $\hat{\mu}_c = \left(e^{1/\lambda}+1-\frac{e^{1/\lambda}-1}{\mu^+}\right)^{-1} \geq 1/2$ we have

$$\mu \geq \widehat{\mu}_c \Leftrightarrow \frac{1}{\mu} \leq e^{1/\lambda}+1-\frac{e^{1/\lambda}-1}{\mu^+} \Leftrightarrow \frac{\mu^+ + 1 - \frac{\mu^+}{\mu}}{1-\mu^+} \geq e^{1/\lambda} \Leftrightarrow \lambda \geq \left(\log\frac{\mu^+ + 1 - \frac{\mu^+}{\mu}}{1-\mu^+}\right)^{-1} = \lambda^*.$$

Proof of Theorem 3 *i*), *ii*), *iii*) and *vi*) follow directly by the optimal error probability functions $\alpha^\star(\mu)$ and $\beta^\star(\mu)$ in Corollary 3 and the fact that $\alpha_m^\star = \mu/\mu^+ \alpha^\star(\mu^+)$ and $\beta_m^\star = \mu/\mu^{+\beta^\star}(\mu^+)$.

iv) Since $\frac{\mu}{\mu^+} < 1$ we have

$$\alpha_m^\star = \frac{\mu}{\mu^+}\alpha^\star(\mu^+) = \frac{\mu\left(e^{1/\lambda}+1\right) - \frac{\mu}{\mu^+}}{e^{2/\lambda}-1} > \frac{\mu\left(e^{1/\lambda}+1\right)-1}{e^{2/\lambda}-1} = \alpha^\star$$

and

$$\beta_m^\star = \frac{\mu}{\mu^+}\beta^\star(\mu^+) = \frac{\frac{\mu}{\mu^+}e^{1/\lambda} - \mu\left(e^{1/\lambda}+1\right)}{e^{2/\lambda}-1} < \frac{e^{1/\lambda} - \mu\left(e^{1/\lambda}+1\right)}{e^{2/\lambda}-1} = \beta^\star.$$

v) $\alpha_m^\star > \alpha^\star$ when

$$\begin{aligned}
\frac{\mu}{\mu^+}\left(1-\mu^+\right) > \frac{\mu\left(e^{1/\lambda}+1\right)-1}{e^{2/\lambda}-1} &\Leftrightarrow \mu\left(\frac{1}{\mu^+}-1\right) > \mu\frac{1}{e^{1/\lambda}-1} - \frac{1}{e^{2/\lambda}-1} \\
&\Leftrightarrow \mu\left(\frac{1}{e^{1/\lambda}-1} - \frac{1}{\mu^+}+1\right) < \frac{1}{e^{2/\lambda}-1} \\
&\Leftrightarrow \mu < \frac{\frac{1}{e^{2/\lambda}-1}}{\frac{1}{e^{1/\lambda}-1} - \frac{1}{\mu^+}+1} = \left(e^{2/\lambda}+e^{1/\lambda}-\frac{e^{2/\lambda}-1}{\mu^+}\right)^{-1} = \widehat{\mu}_{fp}.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\widehat{\mu}_{fp} = \left(e^{2/\lambda}+e^{1/\lambda}-\frac{e^{2/\lambda}-1}{\mu^+}\right)^{-1} &< \left(e^{1/\lambda}+1-\frac{e^{1/\lambda}-1}{\mu^+}\right)^{-1} = \widehat{\mu}_c \\
\Leftrightarrow e^{2/\lambda}-1 > \frac{e^{2/\lambda}-1}{\mu^+} - \frac{e^{1/\lambda}-1}{\mu^+} &\Leftrightarrow e^{1/\lambda}+1 > \frac{e^{1/\lambda}}{\mu^+} \Leftrightarrow \mu^+ > \frac{e^{1/\lambda}}{e^{1/\lambda}+1} = \overline{\mu}
\end{aligned}$$

which is always true by assumption. For the false negative, since $\beta^\star(\mu^+) = 0$, we have $\beta_m^\star = 0 < \beta^\star$.

Proof of Corollary 5 Solving the belief threshold $\hat{\mu}_{fp}$ in Theorem 3 for λ, we obtain the following two roots;

$$\underline{\lambda}_1 = \frac{1}{\log\left[\frac{1}{2}\left(\frac{1}{1-\mu^+} - 1 + \sqrt{\frac{\mu(2-\mu^+)^2 - 4(1-\mu^+)\mu^+}{\mu(1-\mu^+)^2}}\right)\right]}$$
$$\overline{\lambda}_1 = \frac{1}{\log\left[\frac{1}{2}\left(\frac{1}{1-\mu^+} - 1 - \sqrt{\frac{\mu(2-\mu^+)^2 - 4(1-\mu^+)\mu^+}{\mu(1-\mu^+)^2}}\right)\right]}$$

Note that these roots are real valued when expression inside the square root is positive, that is, when $\mu > 4\mu^+ \frac{1-\mu^+}{(2-\mu^+)^2}$. Otherwise there are no real roots and $\alpha_m^\star \geq \alpha^\star$. Assume that this condition holds and consider $\overline{\lambda}_1$. It is positive when

$$\frac{1}{2}\left(\frac{1}{1-\mu^+} - 1 - \sqrt{\frac{\mu(2-\mu^+)^2 - 4(1-\mu^+)\mu^+}{\mu(1-\mu^+)^2}}\right) > 1 \Leftrightarrow \frac{1}{1-\mu^+} - 3 > \sqrt{\frac{\mu(2-\mu^+)^2 - 4(1-\mu^+)\mu^+}{\mu(1-\mu^+)^2}}.$$

Note that first $\mu^+ > 2/3$ should hold so that the left hand side is positive. Then, it can be shown that (after some elementary mathematical operations) $\mu < 1/2$ should hold as well. Similarly, it can be shown that for $\underline{\lambda}_1$ is positive, either when $\mu^+ > 2/3$ or when both $\mu^+ < 2/3$ and $\mu > 1/2$ are satisfied. This means that since $\mu < \mu^+$ by default, when $\mu^+ < 1/2$, there are no positive real-valued roots, and hence $\alpha_m^\star \geq \alpha^\star$. Let us define $\overline{\lambda}_{fp} = \overline{\lambda}_1^+$ and $\underline{\lambda}_{fp} = \underline{\lambda}_1^+$ where $x^+ = \max\{0, x\}$. Assume $\mu^+ > 2/3$. Then when $\mu < 1/2$, $\overline{\lambda}_{fp} = \overline{\lambda}_1$ and $\underline{\lambda}_{fp} = \underline{\lambda}_1$. Taking the first order derivative of the belief threshold $\hat{\mu}_{fp}$ in Theorem 3 with respect to λ, we see that it is positive when $\mu^+ - 2e^{1/x}(1-\mu^+) > 0$, that is, when the two roots $\overline{\lambda}_1$ and $\underline{\lambda}_1$ exist, $\hat{\mu}_{fp}$ is first decreasing than increasing. Then, since $\alpha_m^\star \geq \alpha^\star$ when $\mu \leq \hat{\mu}_{fp}$, it is true also when $\lambda \leq \underline{\lambda}_{fp}$ or $\lambda \geq \overline{\lambda}_{fp}$. Assume now that $1/2 < \mu^+ < 2/3$. Then, when $\mu < 1/2$, $\overline{\lambda}_{fp} = \underline{\lambda}_{fp} = 0$, that is, $\alpha_m^\star \geq \alpha^\star$ for $\lambda > \overline{\lambda}_{fp}$. When $\mu > 1/2$, $\overline{\lambda}_{fp} = 0$ and $\underline{\lambda}_{fp} = \underline{\lambda}_1$, and $\alpha_m^\star \geq \alpha^\star$ for $\lambda < \underline{\lambda}_{fp}$.

Proof of Theorem 4 *i*, *ii*, *iii*, *v* and *vi* correspond to cases where either the DM's prior belief μ or posterior belief μ^+ induces her to spend no cognitive effort. In this case, total cognitive cost is zero in at last one of the cases and the results follow. For case *iv* where the DM processes information in both of these cases, $C_m^\star > C^\star$ when $\frac{\mu}{\mu^+}\left(H\left(\mu^+\right) - \varphi\left(\lambda\right)\right) > H\left(\mu\right) - \varphi\left(\lambda\right)$. Note that the left hand side is a positive increasing function of μ while the right hand side is a concave function that takes its maximum at $\mu = 0.5$. At $\mu = \underline{\mu}$, right hand side is zero and left hand side is positive. At the other extreme when $\mu = \mu^+$, both sides are equal. This means that for $\mu^+ \geq 0.5$, the two functions cross at a single point between $\left(\underline{\mu}, \mu^+\right)$. For $\mu^+ < 0.5$, both functions are increasing. Hence, they cross only if slope of the right hand side function at μ^+ is less than the slope of the left hand side function. The slopes are equal when

$$\frac{H\left(\mu^+\right) - \varphi\left(\lambda\right)}{\mu^+} = \log\frac{1-\mu^+}{\mu^+}\mu^+ \Leftrightarrow \mu^+ = 1 - e^{-\varphi(\lambda)} = \widehat{\mu}_e^+.$$

Hence, for $\mu^+ \leq \widehat{\mu}_e^+ < 0.5$, we have $C_m^\star \geq C^\star$ for all $\mu < \mu^+$. When $\mu^+ \in \left(\widehat{\mu}_e^+, \overline{\mu}\right)$, the unique threshold $\widehat{\mu}_e$ satisfies

$$\frac{\widehat{\mu}_e}{\mu^+}\left(H\left(\mu^+\right) - \varphi\left(\lambda\right)\right) = H\left(\widehat{\mu}_e\right) - \varphi\left(\lambda\right). \tag{A.2}$$

Furthermore, left hand side of $(A.2)$ is decreasing in μ^+ since

$$\frac{H\left(\mu^+\right)}{\mu^+} = \frac{\mu^+\log\frac{1-\mu^+}{\mu^+} - H\left(\mu^+\right)}{\left(\mu^+\right)^2} = \frac{\mu^+\log\frac{1-\mu^+}{\mu^+} + \mu^+\log\mu^+ + \left(1-\mu^+\right)\log\left(1-\mu^+\right)}{\left(\mu^+\right)^2} = \frac{\log\left(1-\mu^+\right)}{\left(\mu^+\right)^2} < 0.$$

Therefore the crossing point that satisfies $(A.2)$ and hence $\widehat{\mu}_e$ is decreasing in μ^+. Lastly, as the concave right hand side function in $(A.2)$ takes its maximum at 0.5, the crossing point is less than that point, i.e., $\widehat{\mu}_e \leq 0.5$.

Proof of Corollary 6 Assume $\mu^+ \geq 0.5$ and $\mu > 1-\mu^+$. Then $H(\mu) > H(\mu^+)$ since H is symmetric around $\mu = 0.5$. Then

$$C_m^\star = \frac{\mu}{\mu^+}\left(H\left(\mu^+\right) - \varphi(\lambda)\right) < \frac{\mu}{\mu^+}\left(H(\mu) - \varphi(\lambda)\right) < H(\mu) - \varphi(\lambda) = C^\star.$$

Assume otherwise. Then $C_m^\star > C^\star$ if $\frac{H(\mu) - \frac{\mu}{\mu^+}H\left(\mu^+\right)}{1-\frac{\mu}{\mu^+}} < \varphi(\lambda)$. We show that the left hand side is increasing in μ. To see this take the first order derivative;

$$\frac{d}{d\mu}\frac{H(\mu) - \frac{\mu}{\mu^+}H(\mu^+)}{\left(1-\frac{\mu}{\mu^+}\right)} = \frac{\left(\log\frac{1-\mu}{\mu} - \frac{H\left(\mu^+\right)}{\mu^+}\right)\left(1-\frac{\mu}{\mu^+}\right) + \frac{1}{\mu^+}\left(H(\mu) - \frac{\mu}{\mu^+}H(\mu^+)\right)}{\left(1-\frac{\mu}{\mu^+}\right)^2}.$$

Simplifying the numerator, we have

$$\begin{aligned}
&\left(1-\frac{\mu}{\mu^+}\right)\log\frac{1-\mu}{\mu} - \frac{H(\mu^+)}{\mu^+} + \frac{H(\mu)}{\mu^+} = \left(\mu^+ - \mu\right)\log\frac{1-\mu}{\mu} + H(\mu) - H\left(\mu^+\right)\\
&= \left(\mu^+ - \mu\right)\log(1-\mu) - \left(\mu^+ - \mu\right)\log\mu - \mu\log\mu - (1-\mu)\log(1-\mu) - H\left(\mu^+\right)\\
&= -\left(1-\mu^+\right)\log(1-\mu) - \mu^+\log\mu - H\left(\mu^+\right).
\end{aligned}$$

This is decreasing in μ as the first order derivative is $\frac{\mu-\mu^+}{1-\mu} < 0$. Evaluating at $\mu = \mu^+$ (which is the largest possible μ) we obtain zero, that is, $-(1-\mu^+)\log(1-\mu^+) - \mu^+\log\mu^+ - H(\mu^+) = 0$. This means the first order derivative of the left hand side is positive. Note also that the right hand side is increasing in λ with $\lim_{\lambda\to\infty}\varphi(\lambda) = \log 2$ while the left hand side is constant. To see this, take the first order derivative $\varphi'(\lambda) = \frac{1}{\lambda^3}\frac{e^{\frac{1}{\lambda}}}{\left(e^{\frac{1}{\lambda}}+1\right)^2} > 0$. Now, when μ is at its maximum, $\mu = \mu^+$, we have

$$\frac{H(1-\mu^+) - \frac{\mu}{\mu^+}H(\mu^+)}{1-\frac{\mu}{\mu^+}} = \frac{\left(1-\frac{\mu}{\mu^+}\right)H(\mu^+)}{1-\frac{\mu}{\mu^+}} = H\left(\mu^+\right) < \varphi(\lambda).$$

The maximum value entropy function H can take is $\log 2$ and for $\mu < \mu^+$, the maximum value that the left hand side can get is less than $\log 2$. Then this means there exists a unique λ that satisfies (11).

Proof of Lemma C.1 We first find the optimal probability $p^\star$ of choosing $a = y$ for the general payoff case. By Theorem 1 in Matějka and McKay (2015), the DM's conditional probability of selecting $a = y$ given $\omega = g$ and $\omega = b$ are respectively, $P_g = \frac{pe^{1/\lambda}}{pe^{1/\lambda}+1-p}$ and $P_b = \frac{pe^{a/\lambda}}{pe^{a/\lambda}+(1-p)e^{c/\lambda}} = \frac{p}{p+(1-p)e^{\delta/\lambda}}$. Her unconditional choice probability p is then

$$p = (1-\mu)\frac{p}{p+(1-p)e^{\delta/\lambda}} + \mu\frac{pe^{1/\lambda}}{pe^{1/\lambda}+1-p}. \tag{A.3}$$

Solving (A.3) yields $\overline{p} = \frac{\mu}{1-e^{-\frac{1}{\lambda}\delta}} - \frac{1-\mu}{e^{\frac{1}{\lambda}}-1}$. Then, similar to the baseline model, $p^\star \leq 0 \Leftrightarrow \mu \leq \underline{\mu}$ and $p^\star \geq 1 \Leftrightarrow \mu \geq \overline{\mu}$ where

$$\underline{\mu} = \frac{1-e^{-\frac{\delta}{\lambda}}}{e^{\frac{1}{\lambda}} - e^{-\frac{\delta}{\lambda}}} \quad \text{and} \quad \overline{\mu} = \frac{e^{\frac{1}{\lambda}}\left(1-e^{-\frac{\delta}{\lambda}}\right)}{e^{\frac{1}{\lambda}} - e^{-\frac{\delta}{\lambda}}}. \tag{A.4}$$

When $\mu \in [\underline{\mu}, \overline{\mu}]$, $p^\star = \overline{p}$. Using Bayes' rule, we have $\gamma(g|y) = p_g \mu / p^\star$ for the posterior belief that the state is good given $a = y$. Plugging in $p^\star$ and p_g, we arrive at $\gamma(g|y)$. Further we have $\gamma(b|y) = 1 - \gamma(g|y)$. The others are found similarly; $\gamma(g|n) = (1 - p_g)\mu / (1 - p^\star)$ and $\gamma(b|n) = 1 - \gamma(g|n)$. Q.E.D.

Writing the decision accuracy in terms of optimal posteriors $A(\mu) = \gamma(b|n)(1 - p^\star) + \gamma(g|y) p^\star$, we see that when $\gamma(b|n) = \gamma(g|y)$, decision accuracy $A(\mu)$ does not depend on prior belief μ. Otherwise, it depends on μ through $p^\star$. By Lemma C.1, $\gamma(b|n) = \gamma(g|y)$ if only if

$$\frac{e^{1/\lambda} - 1}{e^{1/\lambda} - e^{-\delta/\lambda}} = \frac{1 - e^{-\delta/\lambda}}{e^{1/\lambda} - e^{-\delta/\lambda}} e^{1/\lambda} \Leftrightarrow e^{-(\delta-1)/\lambda} = 1$$

which is only possible when $\delta = 1$. Here note that 1 refers to $u(y, g) - u(n, g)$, which is the gain from making the right decision in good state. Therefore when payoff gains across states are equal (i.e., symmetric), accuracy does not depend on prior belief μ.

Proof of Proposition D.1 The accuracy function $A^\star(\mu)$ and value function $V^\star(\mu)$ given in Corollary 2 are convex in μ. Then by Jensen's inequality,

$$A_m^\star = P(X_1 = +) A^\star(\mu_\gamma^+) + P(X_1 = -) A^\star(\mu_\gamma^-) \geq A^\star(\mu) = A^\star$$

since $\mu = P(X_1 = +)\mu_\gamma^+ + P(X_1 = -)\mu_\gamma^-$. The same holds for the expected value.

Proof of Theorem D.1 Here we compare $p^\star = p^\star(\mu)$, $\alpha^\star = \alpha^\star(\mu)$, $\beta^\star = \beta^\star(\mu)$ and $C^\star = C^\star(\mu)$ with $p_m^\star = \frac{\mu}{\mu^+} p^\star(\mu_\gamma^+) + \left(1 - \frac{\mu}{\mu^+}\right) p^\star(\mu_\gamma^-)$, $\alpha_m^\star = \frac{\mu}{\mu^+} \alpha^\star(\mu_\gamma^+) + \left(1 - \frac{\mu}{\mu^+}\right) \alpha^\star(\mu_\gamma^-)$, $\beta_m^\star = \frac{\mu}{\mu^+} \beta^\star(\mu_\gamma^+) + \left(1 - \frac{\mu}{\mu^+}\right) \beta^\star(\mu_\gamma^-)$ and $C_m^\star = \frac{\mu}{\mu^+} C^\star(\mu_\gamma^+) + \left(1 - \frac{\mu}{\mu^+}\right) C^\star(\mu_\gamma^-)$ where $p^\star$ is given in Theorem 1, $C^\star$ in Corollary 2 and $\alpha^\star$ and $\beta^\star$ are given in Corollary 3. Note also that $\mu = \frac{\mu}{\mu^+}\mu_\gamma^+ + \left(1 - \frac{\mu}{\mu^+}\right)\mu_\gamma^-$.

i) (a), (b) and (c) follow since $p^\star(\mu) = \alpha^\star(\mu) = C^\star(\mu) = 0$ for all $\mu \leq \underline{\mu}$ and (c) follows since $\beta^\star$ is linear.

ii) (a), (b) and (d) follow since $p^\star(\mu) = \alpha^\star(\mu) = C^\star(\mu) = 0$ and $p^\star(\mu_\gamma^+) = \alpha^\star(\mu_\gamma^+) = C^\star(\mu_\gamma^+) > 0$. (c) follows by Jensen's inequality since $\beta^*(\mu)$ is concave in $[0, \overline{\mu}]$.

iii) (a) and (b) follow since $p^\star(\mu) = \alpha^\star(\mu) = 0$ and $p^\star(\mu_\gamma^+) = \alpha^\star(\mu_\gamma^+) > 0$. (d) follows since $C^\star(\mu) = C^\star(\mu_\gamma^+) = 0$. (c) follows since $\beta^*(\mu_\gamma^+) = 0$ and $\beta^*(\mu) > \beta^*(\mu_\gamma^-)$.

iv) (a), (b) and (c) follow since $p^*(\mu)$ and $\alpha^\star(\mu)$ are convex and $\beta^*(\mu)$ is concave in $[0, \overline{\mu}]$. (d) Let μ_γ^+ be fixed. Then, $C_m^\star = \frac{\mu}{\mu^+} C^\star(\mu_\gamma^+)$ is linearly increasing in μ from $\frac{\underline{\mu}}{\mu^+} C^\star(\mu_\gamma^+)$ to $C^\star(\mu_\gamma^+)$. Note also that $C^\star(\mu)$ is strictly concave in $[\underline{\mu}, \mu_\gamma^+]$ starting from 0 to $C^\star(\mu_\gamma^+)$. This means two functions cross at a single point and $C_m^\star > C^\star$ when

$$C_m^\star = \frac{\mu}{\mu^+} C^\star(\mu_\gamma^+) > C^\star(\mu) \Longleftrightarrow \varphi(\lambda) > \frac{H(\mu) - \frac{\mu}{\mu^+} H(\mu_\gamma^+)}{1 - \frac{\mu}{\mu^+}}.$$

v) (a), (b), (c) follow since $p^*(\mu)$, $\alpha^\star(\mu)$ and $\beta^*(\mu)$ are linear and $C^\star(\mu)$ is concave in $[\underline{\mu}, \overline{\mu}]$.

vi) (a) $p_m^\star = \frac{\mu}{\mu^+}$ which is linearly increasing in μ. Similarly, $p^*(\mu)$ is linearly increasing in μ in $[\underline{\mu}, \overline{\mu}]$. Note that slope of $\frac{\mu}{\mu^+}$ is lower than $p^*(\mu)$ in $[\underline{\mu}, \overline{\mu}]$ since $p^*(\mu)$ goes from 0 to 1. Then they must cross at a single point. Furthermore, $p_m^\star > p^\star$ when

$$p_m^\star = \frac{\mu}{\mu^+} > \frac{\mu}{1 - e^{-1/\lambda}} - \frac{1 - \mu}{e^{1/\lambda} - 1} \Longleftrightarrow \mu > \left(e^{1/\lambda} + 1 - \frac{e^{1/\lambda} - 1}{\mu^+}\right)^{-1} > \frac{1}{2}.$$

(b) Let μ_γ^+ be fixed. $\alpha^\star\left(\mu_\gamma^-\right)=0$ and $\alpha_m^\star=\frac{\mu}{\mu^+}(1-\mu_\gamma^+)$ is linearly increasing in μ. $\alpha^\star(\mu)$ is also linearly increasing in $\left[\underline{\mu},\overline{\mu}\right]$ with a larger slope (since $\alpha^\star\left(\underline{\mu}\right)=0$ and $\alpha^\star(\overline{\mu})$ is maximum). Then they cross at a single point. Furthermore, $\alpha_m^\star>\alpha^\star$ when

$$\alpha_m^\star=\frac{\mu}{\mu^+}(1-\mu_\gamma^+)>\frac{\mu\left(e^{1/\lambda}+1\right)-1}{e^{2/\lambda}-1}\Longleftrightarrow\mu<\frac{\frac{1}{e^{2/\lambda}-1}}{\frac{1}{e^{1/\lambda}-1}-\frac{1-\mu_\gamma^+}{\mu^+}}.$$

(c) Let μ_γ^+ be fixed. $\beta^\star(\overline{\mu})=0$ and $\beta_m^\star=\left(1-\frac{\mu}{\mu^+}\right)\beta^\star\left(\mu_\gamma^-\right)$ is linearly decreasing in $\mu.\beta^\star(\mu)$ is also linearly decreasing in $\left[\underline{\mu},\overline{\mu}\right]$ with a larger slope (since $\beta^\star\left(\underline{\mu}\right)$ is maximum and $\beta^\star(\overline{\mu})=0$). Then they cross at a single point. Furthermore, $\beta_m^\star>\beta^\star$ when

$$\beta_m^\star=\left(1-\frac{\mu}{\mu^+}\right)\mu_\gamma^->\frac{e^{1/\lambda}-\mu\left(e^{1/\lambda}+1\right)}{e^{2/\lambda}-1}\Longleftrightarrow\mu>\frac{\frac{e^{1/\lambda}}{e^{2/\lambda}-1}-\mu_\gamma^-}{\frac{1}{e^{1/\lambda}-1}-\frac{\mu_\gamma^-}{\mu^+}}.$$

(d) follows since $C^\star(\mu)>0$ and $C^\star\left(\mu_\gamma^+\right)=C^\star\left(\mu_\gamma^-\right)=0$.

vii) (a), (b) and (c) follow since $p^*(\mu)$ and $\alpha^\star(\mu)$ are concave and $\beta^\star(\mu)$ is convex in $\left[\underline{\mu},1\right]$. (d) Let μ_γ^+ be fixed. Then, $C_m^\star=\left(1-\frac{\mu}{\mu^+}\right)C^\star\left(\mu_\gamma^-\right)$ is linearly decreasing in μ from $\frac{\mu}{\mu^+}C^\star\left(\mu_\gamma^-\right)$ to $C^\star\left(\mu_\gamma^-\right)$. Note also that $C^\star(\mu)$ is strictly concave in $\left[\underline{\mu},\overline{\mu}\right]$ starting from $C^\star\left(\mu_\gamma^-\right)$ to 0. This means two functions cross at a single point and $C_m^\star>C^\star$ when

$$C_m^\star=\left(1-\frac{\mu}{\mu^+}\right)C^\star\left(\mu_\gamma^-\right)>C^\star(\mu)\Longleftrightarrow\frac{H(\mu)-\left(1-\frac{\mu}{\mu^+}\right)H\left(\mu_\gamma^-\right)}{\frac{\mu}{\mu^+}}<\varphi(\lambda).$$

viii) (a) follows since $p^\star(\mu)=p^\star(\overline{\mu})=1$ and $p^\star\left(\underline{\mu}\right)=0$ (b) $\alpha^\star\left(\mu_\gamma^-\right)=0$ and for $\mu\geq\overline{\mu}$, $\alpha^\star(\mu)>\alpha^\star\left(\mu_\gamma^+\right)$ as $\alpha^\star(\mu)$ is linearly decreasing in μ. Then $\alpha_m^\star=\frac{\mu}{\mu^+}\alpha^\star\left(\mu_\gamma^+\right)<\alpha^\star(\mu)$ (c) For $\mu\geq\overline{\mu},\beta^\star\left(\mu_\gamma^+\right)=\beta^\star(\mu)=0$. Since $\beta^\star\left(\mu_\gamma^-\right)>0$, $\beta_m^\star>\beta^\star=0$ (d) $C^\star(\mu)=0$ and $C_m^\star$ since $C^\star\left(\mu_\gamma^+\right)=C^\star\left(\mu_\gamma^-\right)=0$.

ix) (a), (b) and (c) follow since $p^*(\mu)$ and $\alpha^\star(\mu)$ are concave and $\beta^\star(\mu)$ is convex in $\left[\underline{\mu},1\right]$. (d) follows since $C^\star\left(\mu_\gamma^+\right)=C^\star(\mu)=0$ and $C^\star\left(\mu_\gamma^-\right)>0$.

x) (a) follows since $p^\star(\mu)=p^\star\left(\mu_\gamma^-\right)=p^\star\left(\mu_\gamma^+\right)=1$. (b) follows since $\alpha^\star(\mu)$ is linear in μ. (c) follows since $\beta^\star(\mu)=\beta^\star\left(\mu_\gamma^-\right)=\beta^\star\left(\mu_\gamma^+\right)=0$. (d) follows since $C^\star(\mu)=C^\star\left(\mu_\gamma^-\right)=C^\star\left(\mu_\gamma^+\right)=0$.

Proof of Proposition D.2 We use Theorem D.1 in Appendix D. Note that $\alpha_m^\star>\alpha^\star$ in cases ii-iv. The union of the regions defined by ii and iii can be represented by $\{\mu<\underline{\mu}\}\&\{\mu_\gamma^+>\underline{\mu}\}$, or equivalently $(\underline{\mu}-\gamma\mu^+)/(1-\gamma)<\mu<\underline{\mu}$ since $\mu_\gamma^+=(1-\gamma)\mu+\gamma\mu^+$. Similarly, case iv and case vi for $\alpha_m^\star>\alpha^\star$ collectively imply the region $\{\mu_\gamma^-<\underline{\mu}\}\&\{\mu_\gamma^+>\underline{\mu}\}\&\{\mu\in[\underline{\mu},\overline{\mu}]\}\&\{\mu<\mu_{fp}^\gamma\}$, or equivalently, $\max\{\underline{\mu},(\underline{\mu}-\gamma\mu^+)/(1-\gamma)\}<\mu<\min\{\underline{\mu}/(1-\gamma),\mu_{fp}^\gamma\}$. These two regions imply then that $\alpha_m^\star>\alpha^\star$ if $(\underline{\mu}-\gamma\mu^+)/(1-\gamma)<\mu<\widehat{\mu}_{fp}^\gamma$ where $\widehat{\mu}_{fp}^\gamma=\min\{\underline{\mu}/(1-\gamma),\mu_{fp}^\gamma\}$ where μ_{fp}^γ is given in Theorem D.1. The same procedure applies for the false negatives.

Proof of Proposition D.3 The conditions in cases ii and iv in Theorem D.1 for $C_m^\star>C^\star$ collectively imply $(\underline{\mu}-\gamma\mu^+)/(1-\gamma)<\mu<\min\{(\overline{\mu}-\gamma\mu^+)/(1-\gamma),\underline{\mu}/(1-\gamma),\mu_e^l\}$. We define $\widehat{\mu}_e^{L\gamma}=\min\{(\overline{\mu}-\gamma\mu^+)/(1-\gamma),\underline{\mu}/(1-\gamma),\mu_e^l\}$. Similarly, cases vii and ix for $C_m^\star>C^\star$ imply $\max\{(\overline{\mu}-\gamma\mu^+)/(1-\gamma),\underline{\mu}/(1-\gamma),\mu_e^h\}<\mu<\overline{\mu}/(1-\gamma)$. We define $\widehat{\mu}^{H\gamma}=\max\{(\overline{\mu}-\gamma\mu^+)/(1-\gamma),\underline{\mu}/(1-\gamma),\mu_e^h\}$.

Proof of Proposition E.1 The consumer problem 1 investigated in Caplin et al. (2019) can be used for our setup as well by taking $M=3$, $u_G=1$, $u_B=0$, $\mu(\omega_k)=\mu_k$ and $\delta=e^{1/\lambda}-1$. Then we apply Theorem 1 in Caplin et al. (2019) to arrive at our characterization.

Proof of Proposition E.2 We first show that the accuracy function in (E.1) in Appendix E can be written as $A^\star(\mu_b,\mu_g)=\max\left\{\mu_1,\frac{e^{1/\lambda}(\mu_1+\mu_2)}{e^{1/\lambda}+1},\frac{e^{1/\lambda}}{e^{1/\lambda}+2}\right\}$ in the domain $\mu_b+\mu_g\in[0,1]$. Assume $\mu_3=1-\mu_1-\mu_2<\frac{1}{e^{1/\lambda}+2}$ and $\mu_1>\mu_2e^{1/\lambda}$. This is equivalent to $\mu_1>\max\left\{\mu_2e^{1/\lambda},\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}-\mu_2\right\}$. Assume $\mu_2>\frac{1}{e^{1/\lambda}+2}$. Then $\mu_1>\mu_2e^{1/\lambda}>\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}-\mu_2$ and since $\mu_2e^{1/\lambda}>\frac{e^{1/\lambda}}{e^{1/\lambda}+2}$, we have $\mu_1>\frac{e^{1/\lambda}}{e^{1/\lambda}+2}$. Assume now that Assume $\mu_2\leq\frac{1}{e^{1/\lambda}+2}$. Then $\mu_1>\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}-\mu_2>\mu_2e^{1/\lambda}$. Since $\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}-\mu_2=\frac{e^{1/\lambda}}{e^{1/\lambda}+2}+\frac{1}{e^{1/\lambda}+2}-\mu_2\geq\frac{e^{1/\lambda}}{e^{1/\lambda}+2}$ we also have $\mu_1>\frac{e^{1/\lambda}}{e^{1/\lambda}+2}$. Note also that $\mu_1>\mu_2e^{1/\lambda}$ implies $\mu_1\left(e^{1/\lambda}+1\right)>e^{1/\lambda}(\mu_1+\mu_2)$, or equivalently $\mu_1>\frac{e^{1/\lambda}(\mu_1+\mu_2)}{e^{1/\lambda}+1}$. Assume now that $\mu_3>\frac{1}{e^{1/\lambda}+2}$. This implies that $1-\mu_1-\mu_2>\frac{1}{e^{1/\lambda}+2}\iff\mu_1+\mu_2<\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}\iff\frac{e^{1/\lambda}}{e^{1/\lambda}+2}>\frac{e^{1/\lambda}(\mu_1+\mu_2)}{e^{1/\lambda}+1}$. When the conditions are reversed, the inequalities are also reversed. This proves that $A^\star(\mu_b,\mu_g)=\max\left\{\mu_1,\frac{e^{1/\lambda}(\mu_1+\mu_2)}{e^{1/\lambda}+1},\frac{e^{1/\lambda}}{e^{1/\lambda}+2}\right\}$ in the domain $\mu_b+\mu_g\in[0,1]$. Then since $\mu_1,\frac{e^{1/\lambda}(\mu_1+\mu_2)}{e^{1/\lambda}+1}$ and $\frac{e^{1/\lambda}}{e^{1/\lambda}+2}$ are all convex functions, then $A^\star(\mu_b,\mu_g)$ is also convex since maximum function is preserved under convexity. As the accuracy function is convex, by Jensen's inequality the machine always increases the DM's expected decision accuracy as $(\mu_b,\mu_g)=P(X_1=+)\left(0,\mu_g^+\right)+P(X_1=-)\left(\mu_b^-,0\right)$ and $A_m^\star=P(X_1=+)A^\star\left(0,\mu_g^+\right)+P(X_1=-)A^\star\left(\mu_b^-,0\right)\geq A^\star(\mu_b,\mu_g)$. A similar approach is also valid for the value function $V^\star$.

Proof of Proposition E.3 Assume $\mu_b=\mu_g=\mu$. We first use Proposition E.1 to find the choice probabilities for the fully symmetric case. First assume that $\mu\leq\frac{1.}{e^{1/\lambda}+2}$. Then since $\mu<1/3$, $\mu_1=1-2\mu$ and $\mu_2=\mu_3=\mu$. Also, since $1-2\mu>\mu e^{1/\lambda}$, by Proposition E.1, $p_1^\star=p_o^\star=1$ and $p_y^\star=p_n^\star=0$. Now assume that $\frac{1.}{e^{1/\lambda}+2}<\mu\leq\frac{1}{3}$. Then, we have $p_y^\star=p_n^\star=\frac{\mu\left(e^{1/\lambda}+2\right)-1}{e^{1/\lambda}-1}$. Now assume $\mu>1/3$. Then $\mu_1=\mu_2=\mu$ and $\mu_3=1-2\mu$. Note also that $\mu_3=1-2\mu>\frac{1}{e^{1/\lambda}+2}$ implies $\mu<\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$. By Proposition E.1, this means for $\frac{1}{3}<\mu<\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$, we have $p_y^\star=p_n^\star=\frac{\mu\left(e^{1/\lambda}+2\right)-1}{e^{1/\lambda}-1}$. Finally, if $\mu\geq\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$, we have $p_y^\star=p_n^\star=\frac{\mu\left(e^{1/\lambda}+2\right)-1}{e^{1/\lambda}-1}$ due to symmetry. Then note that $p_y^\star(\mu,\mu)$ the probability of choosing y can be written as a function of μ as

$$p_y^\star(\mu,\mu)=\begin{cases}0 & \text{if } \mu\leq\frac{1}{e^{1/\lambda}+2}\\ \frac{\mu\left(e^{1/\lambda}+2\right)-1}{e^{1/\lambda}-1} & \text{if } \frac{1}{e^{1/\lambda}+2}<\mu<\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}\\ \frac{1}{2} & \text{if } \mu\geq\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}\end{cases}\tag{A.5}$$

The DM makes a false positive probability when she chooses y when the state is either bad or moderate. In optimality, false positive error probability is $\alpha^\star(\mu,\mu)=p_y^\star(\mu,\mu)\left(1-p_{g|y}^\star\right)$ where $p_{g|y}^\star$ is the posterior probability that the state is good given that the DM chooses y. From the optimal posteriors in Appendix E, we have $p_{g|y}^\star=\frac{e^{1/\lambda}}{e^{1/\lambda}+2}$ if all actions are chosen with positive probability (i.e., when $\frac{1}{e^{1/\lambda}+2}<\mu<\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$), $p_{g|y}^\star=\frac{2\mu e^{1/\lambda}}{e^{1/\lambda}+1}$ if only two options (y and n) are selected (i.e., when $\mu\geq\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$). Finally $p_{g|y}^\star=\mu$ if only one action is selected (i.e., only o is selected which happens when $\mu\leq\frac{1}{e^{1/\lambda}+2}$). Then the false positive rate as a function of μ becomes

$$\alpha^\star(\mu,\mu)=\begin{cases}0 & \text{if } \mu\leq\frac{1}{e^{1/\lambda}+2}\\ 2\frac{\mu\left(e^{1/\lambda}+2\right)-1}{\left(e^{1/\lambda}-1\right)\left(e^{1/\lambda}+2\right)} & \text{if } \frac{1}{e^{1/\lambda}+2}<\mu<\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}\\ \frac{1}{2}-\mu\frac{e^{1/\lambda}}{e^{1/\lambda}+1} & \text{if } \mu\geq\frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}\end{cases}.$$

When the machine provides $x_1=+$ (resp. $x_1=-$), then the DM has only two options good (resp. bad) and moderate with $\mu_g^+=2\mu$ and $\mu_m^+=1-2\mu$ (resp. $\mu_b^-=2\mu$ and $\mu_m^-=1-2\mu$). This means the DM can only make false positive errors when $x_1 = +$ which happens with probability 1/2. To find the false positive error rate in this case, we use the characterization in Corollary 3. Note that both error characterizations have two belief thresholds. Then we have the following different cases:

- If $\mu \le \frac{1}{2}\frac{1}{e^{1/\lambda}+1}$, then $2\mu < \frac{1}{e^{1/\lambda}+1}$ and $\alpha_m^\star = \alpha^\star = 0$.
- If $\frac{1}{2}\frac{1}{e^{1/\lambda}+1} < \mu \le \frac{1}{e^{1/\lambda}+2}$, then $2\mu > \frac{1}{e^{1/\lambda}+1}$ and hence $\alpha_m^\star > \alpha^\star = 0$. Note also that $\frac{1}{e^{1/\lambda}+2} > \frac{1}{2}\frac{1}{e^{1/\lambda}+1}$ for each $\lambda > 0$.
- If $\frac{1}{e^{1/\lambda}+2} < \mu \le \frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1}$ (or if $\frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1} \le \mu < \frac{1}{e^{1/\lambda}+2}$ depending on which one is bigger) then $\alpha_m^\star > \alpha^\star$ when $\frac{1}{2}\frac{2\mu\left(e^{1/\lambda}+1\right)-1}{e^{2/\lambda}-1} > 2\frac{\mu\left(e^{1/\lambda}+2\right)-1}{\left(e^{1/\lambda}-1\right)\left(e^{1/\lambda}+2\right)}$ which (after some mathematical manipulation) reduces to $\mu < \frac{\frac{3}{2}e^{1/\lambda}+1}{\left(e^{1/\lambda}+1\right)\left(e^{1/\lambda}+2\right)}$.However, this is only valid when $\lambda < 1/\ln 2$ since otherwise the threshold is greater than $\frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1}$.
- If $\frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1} < \mu < \frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$, then, $\alpha_m^\star > \alpha^\star$ if $\frac{1}{2}(1-2\mu) > 2\frac{\mu\left(e^{1/\lambda}+2\right)-1}{\left(e^{1/\lambda}-1\right)\left(e^{1/\lambda}+2\right)}$which reduces to $\mu < \frac{\frac{1}{2}\left(e^{1/\lambda}-1\right)\left(e^{1/\lambda}+2\right)+2}{\left(e^{1/\lambda}+2\right)\left(e^{1/\lambda}+1\right)}$. The threshold is only valid when $\lambda > 1/\ln 2$, since otherwise it is less than $\frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1}$.
- If $\mu \ge \frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$, then $\alpha_m^\star < \alpha^\star$ since $\frac{1}{2}(1-2\mu) < \frac{1}{2} - \mu\frac{e^{1/\lambda}}{e^{1/\lambda}+1}$ for each $\lambda > 0$.

Taken together, $\alpha_m^\star > \alpha^\star$ only if $\frac{1}{2}\frac{1}{e^{1/\lambda}+1} < \mu < \min\left\{\frac{\frac{3}{2}e^{1/\lambda}+1}{\left(e^{1/\lambda}+1\right)\left(e^{1/\lambda}+2\right)}, \frac{\frac{1}{2}\left(e^{1/\lambda}-1\right)\left(e^{1/\lambda}+2\right)+2}{\left(e^{1/\lambda}+2\right)\left(e^{1/\lambda}+1\right)}\right\} = \mu_{fp}^*$. Note that the first component is increasing in λ while the second one is decreasing and it can be verified that they intersect at $\lambda = 1.4427$ which gives a value of 1/3. This means the threshold cannot be greater than 1/3 due to the minimum function.

Proof of Proposition E.4 Using the characterization for the general case in Appendix E, the cognitive effort function for the symmetric case can be written as

$$C^\star(\mu,\mu) = \begin{cases} 0 & \text{if } \mu \le \frac{1}{e^{1/\lambda}+2} \\ \lambda\left[H(\mu,\mu) - \varphi^{(2)}(\lambda)\right] & \frac{1}{e^{1/\lambda}+2} < \mu < \frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2} \\ 2\lambda\mu\left[\ln 2 - \varphi(\lambda)\right] & \mu \ge \frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2} \end{cases}$$

where $\varphi^{()}(\lambda)$ and $\varphi(\lambda)^{(2)}$ are given in (E.3) and (E.4) in Appendix E. Finally, $H(\mu,\mu) = -2\mu\ln(\mu) - (1-2\mu)\ln(1-2\mu)$. We now compare the DM's cognitive effort with and without the machine. Note that when the machine gives any information $x_1 = +$ or $x_1 = -$ (which happens with probability 1/2 in the fully symmetric case), the DM has moderate state and good or bad state to consider respectively, which is our base model. Since the DM's posterior beliefs are equal at 2μ for either case and since the effort function in our base case is symmetric, the DM's expected cognitive effort in the fully symmetric case is $C_m^\star = C^\star(2\mu)$ in our base case (see Corollary 2). Then we can use an approach similar to the Proof of Proposition E.3 . We have the following cases:

- If $\mu \le \frac{1}{2}\frac{1}{e^{1/\lambda}+1}$, then $C_m^\star = C^\star = 0$ since the DM does not process information with or without the machine.
- If $\frac{1}{2}\frac{1}{e^{1/\lambda}+1} < \mu \le \frac{1}{e^{1/\lambda}+2}$, then $C_m^\star > C^\star = 0$ since the DM does not process information without the machine, but the machine induces her to process information.
- If $\frac{1}{e^{1/\lambda}+2} < \mu \le \frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1}$, then $C_m^\star > C^\star$ if

$$H(2\mu,0) - \varphi(\lambda) > H(\mu,\mu) - \varphi^{(3)}(\lambda)$$
$$-2\mu\ln(2\mu) - (1-2\mu)\ln(1-2\mu) - \varphi(\lambda) > -2\mu\ln(\mu) - (1-2\mu)\ln(1-2\mu) - \varphi^{(3)}(\lambda)$$
$$\mu < \frac{\varphi^{(3)}(\lambda) - \varphi(\lambda)}{2\ln 2}.$$

Depending on the level of λ, this threshold may be greater than the end point $\frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1}$. In that case, $C_m^\star > C^\star$ in the whole region.

- If $\frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1} < \mu < \frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$, then $0 = C_m^\star < C^\star$ since the DM without the machine does not process information as $2\mu > \frac{e^{1/\lambda}}{e^{1/\lambda}+1}$.
- If $\mu \geq \frac{1}{2}\frac{e^{1/\lambda}+1}{e^{1/\lambda}+2}$, then $0 = C_m^\star < C^\star$ as well. All together, $C_m^\star > C^\star$ if and only if $\frac{1}{2}\frac{1}{e^{1/\lambda}+1} < \mu < \mu_e^*$ where $\mu_e^* = \min\left\{\frac{1}{2}\frac{e^{1/\lambda}}{e^{1/\lambda}+1}, \frac{\varphi^{(3)}(\lambda)-\varphi(\lambda)}{2\ln 2}\right\}$. Finally, note that the second component in μ_e^* is increasing and in the limit equal to $\ln(3/2)/\ln(4) < 1/3$.

Appendix B: General Payoff Structure

Our base model assumes that the DM's payoff corresponds to the overall accuracy of her decisions. Our framework can also account for a general payoff structure of the form $u(a,\omega)$, for $(a,\omega) \in \{y,n\} \times \{g,b\}$. More specifically, as we show below, we can normalize any payoff structure $u(a,\omega)$ such that $u(y,g) = 1$ and $u(n,g) = 0$ without loss of generality. To avoid any trivial solution, we assume that $u(n,b) > u(y,b)$ (otherwise, the payoff of $a = y$ dominates the payoff of $a = n$ in all states of the world and the DM directly chooses the former without processing information). In this setup, difference $\delta = u(n,b) - u(y,b)$ denotes the net value of correctly identifying the bad state. (The net value of correctly identifying the good state is always equal to one.) Thus, the DM prefers to correctly identify the bad state over the good state if and only if $\delta > 1$. In our base model, $\delta = 1$ with $u(n,b) = 1$ and $u(y,b) = 0$, so that the DM is indifferent between identifying the good and the bad states.

B.1. Normalizing the Payoffs

One can transform any general payoff matrix $\widehat{u}(a,\omega)$ with $a \in \{y,n\}$ and $\omega \in \{g,b\}$ by first subtracting $\widehat{u}(n,g)$ from each payoff, and then scaling each by $1/(\widehat{u}(y,g) - \widehat{u}(n,g))$. Then, the new payoff structure becomes

$$u(n,g) = \widehat{u}(n,g) - \widehat{u}(n,g) = 0 \qquad u(y,g) = \frac{\widehat{u}(y,g) - \widehat{u}(n,g)}{\widehat{u}(y,g) - \widehat{u}(n,g)} = 1$$
$$u(y,b) = \frac{\widehat{u}(y,b) - \widehat{u}(n,g)}{\widehat{u}(y,g) - \widehat{u}(n,g)} = a \qquad u(n,b) = \frac{\widehat{u}(n,b) - \widehat{u}(n,g)}{\widehat{u}(y,g) - \widehat{u}(n,g)} = c.$$

Information cost parameter λ, should then be scaled by $1/(u(y,g) - u(n,g))$, to arrive at an identical behavioral structure. That is, the new information cost should be $\lambda' = \frac{\lambda}{u(y,g)-u(n,g)}$. The reason is that subtracting $\widehat{u}(n,g)$ from each payoff does not change the DM's problem since the payoff differences (i.e., incentives) stay the same. Therefore, there is no need to change λ. However scaling each payoff by a constant also scales the differences between them which creates a different incentive structure. To avoid this, one needs to scale the information cost also by the same constant.

B.2. Impact of the Machine for General Payoffs

Figure B.1 depicts the impact of the machine on the DM's decision (analogous to Figure 3) for $\delta < 1$ and $\delta > 1$. The figures demonstrate that the structure of our result continues to hold for more general payoffs. In addition, the figure reveals that the set of values of beliefs μ and μ^+ for which $p_m^\star \geq p^\star$ widens as δ increases. Indeed, increasing δ decreases the likelihood that the DM will choose $a = y$ as this option becomes a less attractive alternative. Accordingly, the threshold level $\overline{\mu}$ on the DM's prior belief that warrants immediate

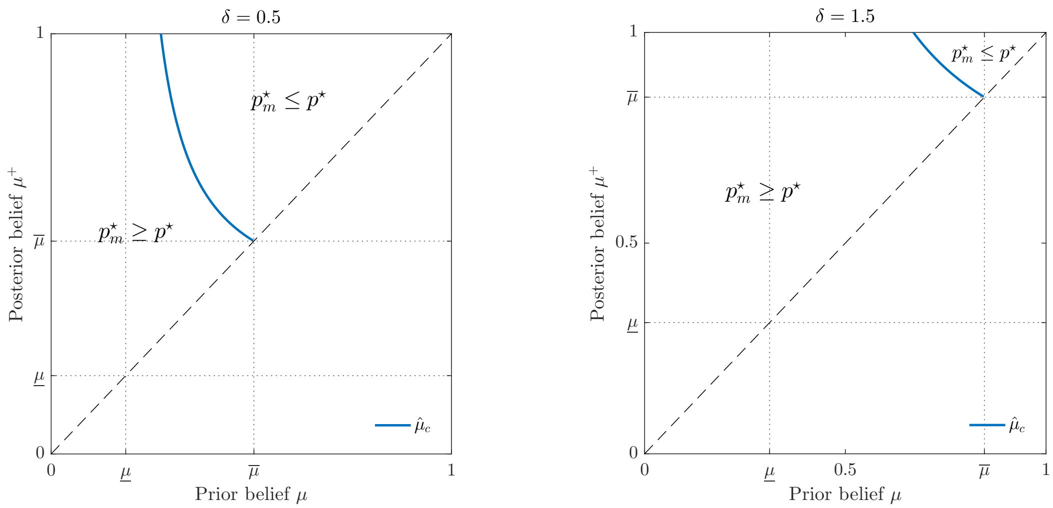


Figure B.1 Impact of incentive structures on DM's decision ($\lambda = 1$)

ex-ante $a = y$ decision increases. That is, the DM needs to be more confident about the good state to choose $a = y$ without the need to spend further cognitive effort. According to Theorem 2, we already know that the machine induces the DM to choose $a = y$ when her posterior is less than $\overline{\mu}$.

When DM's incentives change, the machine's impact on the extent of errors that the DM makes and the expected cognitive effort do not structurally change. In particular, as in our baseline model, when the machine assists the DM with some accurate information, the DM's false negative error always decreases as it completely eliminates the possibility of bad state in some cases. Similarly, the machine can increase the DM's propensity to make false positive errors in some cases. In particular, there still exists a unique threshold $\hat{\mu}_{fp}$ on the DM's prior belief that determines whether the DM makes more or fewer false positive errors with the machine. Furthermore, the larger the net value of correctly identifying bad state δ, the larger the parameter space where the DM makes more false positive errors with the machine. This is because the region where the DM is inclined to choose $a = y$ more with the machine is larger (see Figure B.1). The effect of δ on DM's propensity to make false a positive error is illustrated in Figure B.2.

Changing incentives also has a significant effect on the amount of cognitive cost that the DM incurs. In particular, the more at stake, the more cognitive effort the DM tolerates expending. As in our baseline case, the machine can only increase the DM's ex-ante effort when both her prior and posterior with the machine-supplied information induce the DM to exert cognitive effort (i.e., $\underline{\mu}, \mu < \mu^+, \overline{\mu}$). Therefore, as δ increases, the parameter region where the DM induces the DM to exert more cognitive effort increases as the difference $\overline{\mu} - \underline{\mu}$ becomes larger. This is illustrated in Figure B.3.

Appendix C: Invariance of Accuracy to Prior Belief

We show this property by writing the DM's decision accuracy in terms of the optimal posteriors the DM constructs. The following lemma gives the characterization of these posteriors in the general payoff case.

Lemma C.1. *DMs optimal posterior beliefs when $\mu \in \left(\underline{\mu}, \overline{\mu}\right)$ are*

$$\gamma\left(g|n\right) = \frac{1 - e^{-\delta/\lambda}}{e^{1/\lambda} - e^{-\delta/\lambda}}$$

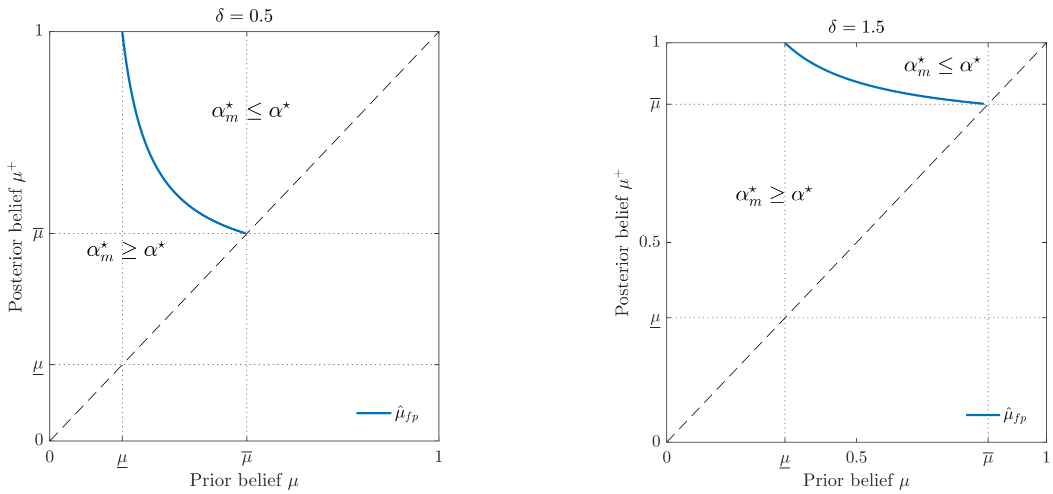


Figure B.2 Impact of incentive structures on DM's false positive error rate ($\lambda = 1$)

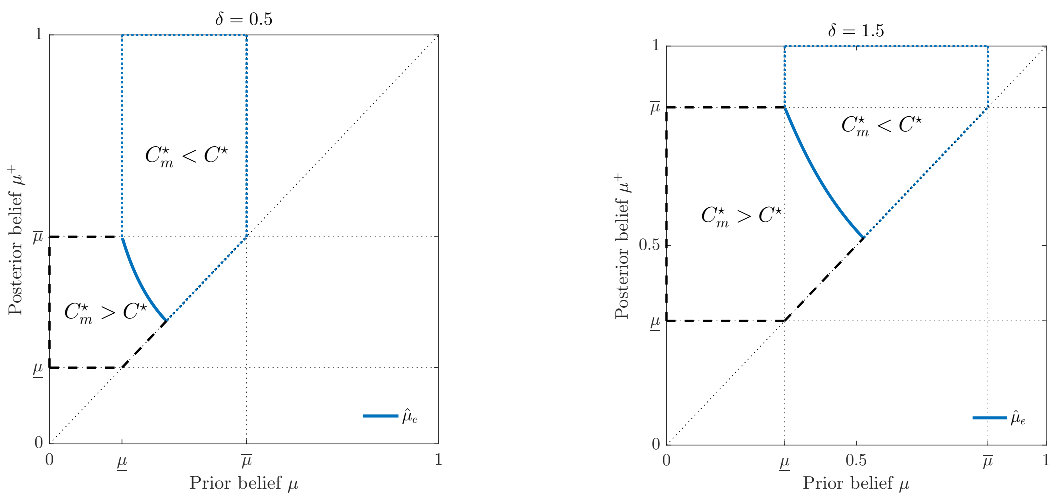


Figure B.3 Impact of incentive structures on DM's cognitive effort ($\lambda = 1$)

$$\gamma\left(g|y\right) = \frac{1 - e^{-\delta/\lambda}}{e^{1/\lambda} - e^{-\delta/\lambda}} e^{1/\lambda}$$

with $\gamma(b|y) = 1 - \gamma(g|y)$ *and* $\gamma(b|n) = 1 - \gamma(g|n)$.

Appendix D: Incorporating Trust

In this section we extend our baseline model to account for possible biases that the DM may hold against (or toward) the machine. The mistrust in machine implies that the DM may not fully believe that the state is bad when the machine's signal is negative. Hence, the machine is not seen as perfectly accurate. Indeed, as we explain at the end of section, our proposed framework also captures the case of an imprecise machine that may generate false positive or negative errors.

We follow the behavioral operations literature (e.g., Özer et al. 2011) in modeling DM's trust and bias towards the machine. Specifically, we assume that given machine input $x_1 \in \{+, -\}$, the DM updates her belief according to $\mu_\gamma^+ = (1-\gamma)\,\mu + \gamma\mu^+$ and $\mu_\gamma^- = (1-\gamma)\,\mu$, where higher values of trust parameter $\gamma \in [0,1]$ indicates more trust in the machine. In other words, the DM mixes her prior belief μ with the posterior belief she would have were she to fully trust the machine. Note that $\gamma = 1$ retrieves our base model, while the DM fully ignores the machine and always decides alone when $\gamma = 0$. For $0 < \gamma < 1$, the DM's level of trust weakens the effect of the machine input on the DM's belief, i.e. $\mu^- = 0 < \mu_\gamma^- < \mu < \mu_\gamma^+ < \mu^+$. In particular, the negative signal of the machine does not fully reveal the bad state, that is $\mu_\gamma^- > 0$ in this case. Nonetheless, we show next that the machine always improves the DM's expected accuracy and value for any trust level.

PROPOSITION D.1. *For any given $\lambda > 0$ and $\gamma \in [0,1]$, we have $A_m^\star \geq A^\star$ and $V_m^\star \geq V^\star$.*

To investigate the impact of the machine on the DM's behavior in this generalized setup, we follow the approach of Section 5 and first extend Theorem 1. This allows generalizing Theorems 2, 3, and 4 for any trust parameter $\gamma \in [0,1]$. Note that in addition to prior belief μ and posterior beliefs μ_γ^+, our results also depend on $\mu_\gamma^- > 0$ when $\gamma \in (0,1)$. Because of this, the number of parameter regions that describe the effect of the machine increases from six when $\mu^- = 0$, as in the base case, to ten when $\mu_\gamma^- > 0$. Theorem D.1 provides a full characterization of the impact of the machine on the DM's choice probability, false positive/negative error rates and cognitive effort. Note that we assume $\mu_\gamma^- < \mu < \mu_\gamma^+$.

THEOREM D.1. *Given information cost $\lambda > 0$ and $\mu < \mu^+$, we have*

i) *If $\mu_\gamma^+ \leq \underline{\mu}$, then, (a) $p_m^\star = p^\star = 0$. (b) $\alpha_m^\star = \alpha^\star = 0$. (c) $\beta_m^\star = \beta^\star > 0$. (d) $C_m^\star = C^\star = 0$.*

ii) *If $\mu \leq \underline{\mu}$ and $\mu_\gamma^+ \in \left(\underline{\mu}, \overline{\mu}\right)$, then, (a) $p_m^\star > p^\star = 0$. (b) $\alpha_m^\star > \alpha^\star = 0$. (c) $\beta_m^\star < \beta^\star$. (d) $C_m^\star > C^\star = 0$.*

iii) *If $\mu \leq \underline{\mu}$ and $\mu_\gamma^+ \geq \overline{\mu}$, then (a) $p_m^\star > p^\star = 0$. (b) $\alpha_m^\star > \alpha^\star = 0$. (c) $\beta_m^\star < \beta^\star$. (d) $C_m^\star = C^\star = 0$.*

iv) *If $\mu_\gamma^- \leq \underline{\mu}$ and $\mu, \mu_\gamma^+ \in \left[\underline{\mu}, \overline{\mu}\right]$, then (a) $p_m^\star > p^\star$. (b) $\alpha_m^\star > \alpha^\star$. (c) $\beta_m^\star < \beta^\star$.*

(d) threshold $\widehat{\mu}_e^l$ exists such that $C_m^\star > C^\star$ if $\mu < \widehat{\mu}_e^l$ and $C_m^\star \leq C^\star$ otherwise. Furthermore, threshold $\widehat{\mu}_e^l$ uniquely solves (for μ)

$$\frac{H(\mu) - \frac{\mu}{\mu^+} H\left(\mu_\gamma^+\right)}{1 - \frac{\mu}{\mu^+}} = \varphi(\lambda)$$

v) *If $\mu_\gamma^-, \mu, \mu_\gamma^+ \in \left[\underline{\mu}, \overline{\mu}\right]$, then, (a) $p_m^\star = p^\star$. (b) $\alpha_m^\star = \alpha^\star$. (c) $\beta_m^\star = \beta^\star$. (d) $C_m^\star < C^\star$*

vi) *If $\mu_\gamma^- \leq \underline{\mu}$ and $\mu \in \left[\underline{\mu}, \overline{\mu}\right]$ and $\mu_\gamma^+ \geq \overline{\mu}$, then,*

(a) *threshold $\widehat{\mu}_c$ exists such that $p_m^\star > p^\star$ if $\mu < \widehat{\mu}_c$ and $p_m^\star \leq p^\star$ otherwise. Furthermore, threshold $\widehat{\mu}_c$ is given as*

$$\widehat{\mu}_c = \left(e^{1/\lambda} + 1 - \frac{e^{1/\lambda} - 1}{\mu^+}\right)^{-1} \geq \frac{1}{2}.$$

(b) *threshold $\widehat{\mu}_{fp}^\gamma$ exists such that $\alpha_m^\star > \alpha^\star$ if $\mu < \widehat{\mu}_{fp}^\gamma$ and $\alpha_m^\star \leq \alpha^\star$ otherwise. Furthermore, the threshold $\widehat{\mu}_{fp}^\gamma$ uniquely solves (for μ)*

$$\frac{\mu\left(e^{1/\lambda} + 1\right) - 1}{e^{2/\lambda} - 1} = \frac{\mu}{\mu^+}\left(1 - \mu_\gamma^+\right)$$

(c) *threshold $\widehat{\mu}_{fn}^\gamma$ exists such that $\beta_m^\star < \beta^\star$ if $\mu < \widehat{\mu}_{fn}^\gamma$ and $\beta_m^\star \geq \beta^\star$ otherwise. Furthermore, the threshold $\widehat{\mu}_{fn}^\gamma$ uniquely solves (for μ)*

$$\frac{e^{1/\lambda} - \mu\left(e^{1/\lambda} + 1\right)}{e^{2/\lambda} - 1} = \left(1 - \frac{\mu}{\mu^+}\right)\mu_\gamma^-$$

(d) $0 = C_m^\star < C^\star$

vii) *If $\mu_\gamma^-, \mu \in \left[\underline{\mu}, \overline{\mu}\right]$ and $\mu_\gamma^+ \geq \overline{\mu}$, then, (a) $p_m^\star < p^\star$. (b) $\alpha_m^\star < \alpha^\star$. (c) $\beta_m^\star > \beta^\star$.*

(d) threshold $\widehat{\mu}_e^h$ exists such that $C_m^\star < C^\star$ if $\mu < \widehat{\mu}_e^h$ and $C_m^\star \geq C^\star$ otherwise. Furthermore, threshold $\widehat{\mu}_e^h$ uniquely solves (for μ)

$$\frac{H(\mu) - \left(1 - \frac{\mu}{\mu^+}\right) H\left(\mu_\gamma^-\right)}{\frac{\mu}{\mu^+}} = \varphi(\lambda)$$

viii) If $\mu_\gamma^- \leq \underline{\mu}$ and $\mu_\gamma^+\, \mu \geq \overline{\mu}$, then, (a) $p_m^\star < p^\star = 1$. (b) $\alpha_m^\star < \alpha^\star$. (c) $\beta_m^\star > \beta^\star$. (d) $C_m^\star = C^\star = 0$.

ix) If $\mu_\gamma^- \in \left[\underline{\mu}, \overline{\mu}\right]$ and $\mu_\gamma^+, \mu \geq \overline{\mu}$, then, (a) $p_m^\star < p^\star = 1$. (b) $\alpha_m^\star < \alpha^\star$. (c) $\beta_m^\star > \beta^\star = 0$. (d) $C_m^\star > C^\star = 0$

x) If $\mu_\gamma^- \geq \overline{\mu}$, then (a) $p_m^\star = p^\star = 1$. (b) $\alpha_m^\star = \alpha^\star$. (c) $\beta_m^\star = \beta^\star = 0$. (d) $C_m^\star = C^\star = 0$.

The following results highlight how trust interacts with the effect of the machine on the DM's decision making process.

PROPOSITION D.2. *For any given $\lambda > 0$, $\gamma \in (0,1]$ and $\mu^+ > \mu$, thresholds $\widehat{\mu}_{fp}^\gamma$ and $\widehat{\mu}_{fn}^\gamma$ exist such that*

i) $\alpha_m^\star > \alpha^\star$ if and only if $\frac{\underline{\mu} - \gamma\mu^+}{1-\gamma} < \mu < \widehat{\mu}_{fp}^\gamma$,

ii) $\beta_m^\star > \beta^\star$ if and only if $\widehat{\mu}_{fn}^\gamma < \mu < \frac{\overline{\mu}}{1-\gamma}$.

Proposition D.2 shows that the machine may continue to increase the DM's propensity of making false positive errors in the presence of mistrust. As in our base model, this happens when the DM does not strongly favor the good state a priori (i.e., $\mu < \widehat{\mu}_{fp}^\gamma$). In this sense, the result in Theorem 3 is robust to the inclusion of trust. In contrast to our base model, however, the machine may also increase the DM's propensity to make false negative errors. This is due to the DM's mistrust in the machine's negative signal, which yields $\mu_\gamma^- > 0$. This happens when the DM strongly favors the good state a priori (i.e., $\mu > \widehat{\mu}_{fn}^\gamma$).

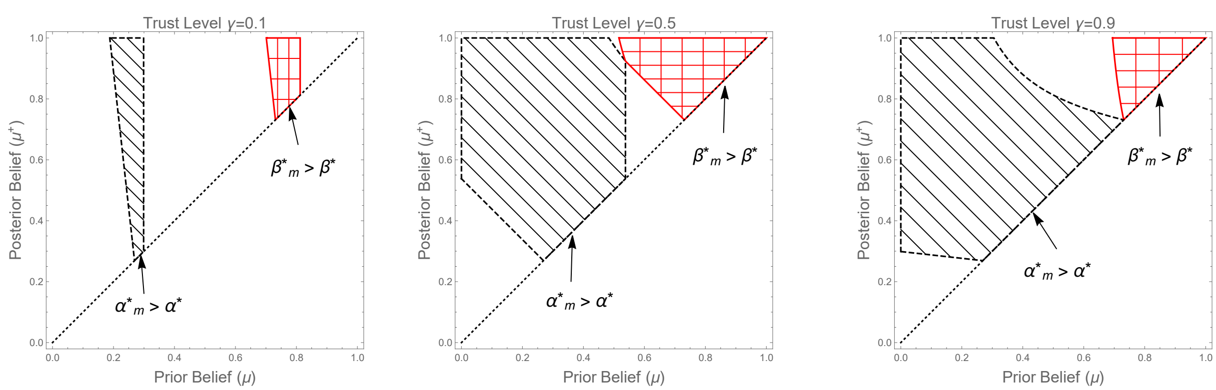


Figure D.1 Impact of the machine on the DM's decision errors for different trust levels

Figure D.1 illustrates the parameter regions where the machine increases the DM's errors for three different trust levels. Since the machine always increases DM's decision accuracy as per Proposition D.1, the machine never increases both false positive and negative rates at the same time.

The DM's bias toward the machine also interacts with the effect the machine has on the DM's cognitive effort, as we show next.

PROPOSITION D.3. *For any given $\lambda > 0$, $\gamma \in (0,1]$ and $\mu^+ > \mu$, thresholds $\widehat{\mu}_e^{L\gamma}$ and $\widehat{\mu}_e^{H\gamma}$ exist such that $C_m^\star > C^\star$ if and only if $\frac{\underline{\mu} - \gamma\mu^+}{1-\gamma} < \mu < \widehat{\mu}_e^{L\gamma}$ or $\widehat{\mu}_e^{H\gamma} < \mu < \frac{\overline{\mu}}{1-\gamma}$.*

Proposition D.3 shows that the machine may increase the DM's cognitive effort in this setup as well. This can happen when the DM sufficiently favors either the bad state (i.e., $\mu < \widehat{\mu}_e^{L\gamma}$), or the good state (i.e., $\mu > \widehat{\mu}_e^{H\gamma}$). The former case is consistent with Theorem 4 in the base model, and a similar rationale holds. The second case, however, does not occur in our base model. This is because the DM's posterior belief upon a negative machine signal is positive if $\gamma > 0$. This may then correspond to an increase in task difficulty and thus induces the DM to process more information.

Figure D.2 illustrates the parameter regions in which the machine increases the DM's cognitive effort for different values of γ. The figure depicts two such regions, one for each of the two intervals in μ defined by Proposition D.3. Note that when trust level γ is sufficiently high, the region corresponding to the second interval disappears since posterior μ_γ^- approaches to zero, which corresponds to our base model.

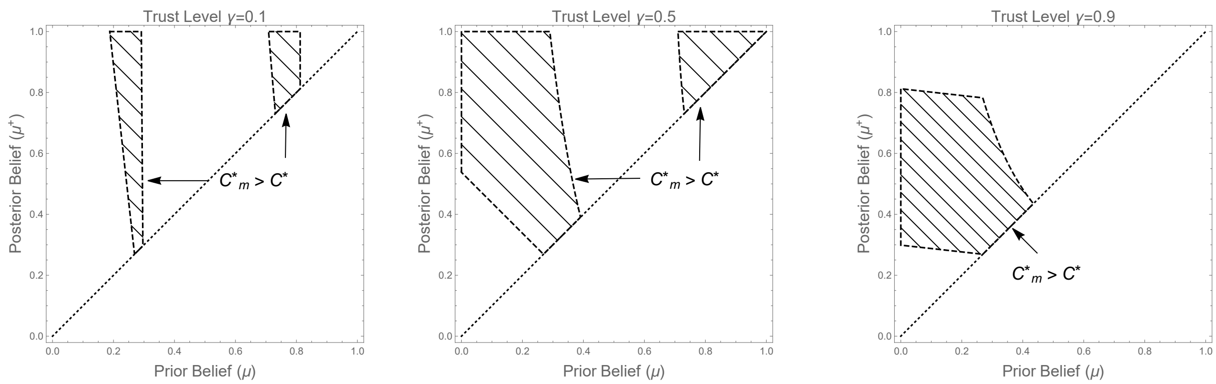


Figure D.2 **Impact of machine on DM's cognitive effort for different trust levels**

A key point in the above analysis is that the DM's optimal choice probabilities, decision errors and cognitive efforts only depend on prior belief μ (without the machine) and posterior beliefs μ_γ^+ and μ_γ^- (with the machine) in addition to λ. In our original setup $\mu_\gamma^- = 0$, but Theorem 5 extends these results to a setup where μ, μ_γ^+ and μ_γ^- are all free parameters, with $\mu_\gamma^- \geq 0$. Importantly, this characterization is context free, i.e., once these probabilities are specified, the previous approach holds. (Theorem D.1 also provides thresholds $\widehat{\mu}_{fp}^\gamma$, $\widehat{\mu}_{fn}^\gamma$, $\widehat{\mu}_e^{L\gamma}$ and $\widehat{\mu}_e^{H\gamma}$ of Propositions D.2 and D.3 in closed form.).

D.1. Incorporating Machine Imprecision

Assume that the input the machine provides is not always accurate. In particular, assume that the machine gives signal $Y \in \{+, -\}$ with false positive error probability $\gamma_1 = P(Y = + | \omega = b)$ and false negative error probability $\gamma_2 = P(Y = - | \omega = g)$. Upon receiving a positive signal from the machine, the DM updates her prior as

$$\mu_\gamma^+ = P(\omega = g \mid Y = +) = \frac{(1-\gamma_2)\,\mu}{(1-\gamma_2)\,\mu + \gamma_1\,(1-\mu)}$$
$$\mu_\gamma^- = P(\omega = g \mid Y = -) = \frac{\gamma_2\mu}{\gamma_2\mu + (1-\gamma_1)\,(1-\mu)}.$$

Assuming $\gamma_1 + \gamma_2 < 1$ ensures that $\mu_\gamma^- < \mu < \mu_\gamma^+$ and Theorem D.1 can directly be used to assess the impact of the machine on the DM's choice behavior. If $\gamma_1 + \gamma_2 > 1$, then $\mu_\gamma^+ < \mu < \mu_\gamma^-$ and it just enough to relabel μ_γ^+ and μ_γ^- and apply Theorem D.1. Note that it is possible to combine mistrust and imprecision to model situations where the DM is biased *and* the machine is inaccurate.

Appendix E: A Symmetric Setting with an Additional State

In our base model, the machine reduces the DM's uncertainty in an asymmetric way as it fully resolves the bad state for the DM when the first information source X_1 is negative. In order to account for a symmetric reduction setting, we introduce a third state, which we call *moderate* (denoted by $\omega = m$) and a corresponding accurate decision, which is declaring the test as *inconclusive* (denoted by $a = o$). We assume that the true state is *good* (resp. *bad*) if and only if both X_1 and X_2 are positive (resp. negative). Otherwise (i.e., if $(X_1, X_2) \in \{(+,-),(-,+)\}$), the state is assumed to be *moderate*. In this setup, the machine never fully resolves the DM's uncertainty. That is, although a negative signal rules out the good state, the DM may still need to process information to distinguish the bad from the moderate state.

It is well-known (see Caplin et al. 2019) that with more than two states, the DM may naturally rule out some of the states a priori before eliciting any signal (i.e., forms *considerations sets*). To accommodate for this extension, we slightly change our notation and let $p^\star_a(\mu_b, \mu_g)$ denote the unconditional probability that the DM chooses $a \in \{n, y, o\}$ as a function her prior belief $\mu_b = \pi(-,-)$ that the state is bad and $\mu_g = \pi(+,+)$ that the state is good. The DM's prior belief for the moderate state is then $\mu_m = 1 - \mu_b - \mu_g$.

We characterize next the DM's choice probabilities as a function of her prior belief for a given information cost λ. To that end, we order the priors such that $\mu_1 \equiv \max(\mu_b, \mu_g, \mu_o)$, $\mu_3 \equiv \min(\mu_b, \mu_g, \mu_o)$ and μ_2 denotes the remaining prior with $\mu_1 \geq \mu_2 \geq \mu_3$. Action $a_i, i \in \{1,2,3\}$ indicates then the action corresponding to the state associated with prior μ_i. For instance, if $\mu_1 = \mu_g$, $\mu_2 = \mu_b$ and $\mu_3 = \mu_m$ then $a_1 = y$, $a_2 = n$ and $a_3 = o$. We also have $\mu_m = 1 - \mu_b - \mu_g$ and thus present our result in the parameter space $\mu_g \times \mu_b$.

Proposition E.1. *Let $\mu_3 = \min\{\mu_g, \mu_b, 1 - \mu_b - \mu_g\}$, $\mu_1 = \max\{\mu_g, \mu_b, 1 - \mu_b - \mu_g\}$, and $\mu_2 = 1 - \mu_1 - \mu_3$. Let a_1, a_2 and a_3 denote the corresponding correct actions. We have*
- *If $\mu_3 > \frac{1}{e^{1/\lambda}+2}$, then $p^\star_{a_i}(\mu_b, \mu_g) = \frac{\mu_i\left(e^{1/\lambda}+2\right)-1}{e^{1/\lambda}-1}$ for $i \in \{1,2,3\}$.*
- *If $\mu_3 < \frac{1}{e^{1/\lambda}+2} \& \mu_1 < \mu_2 e^{1/\lambda}$, then $p^\star_{a_i}(\mu_b, \mu_g) = \frac{\frac{\mu_i}{\mu_1+\mu_2}\left(e^{1/\lambda}+1\right)-1}{e^{1/\lambda}-1}$ for $i \in \{1,2\}$ & $p^\star_{a_3}(\mu_b, \mu_g) = 0$.*
- *If $\mu_3 < \frac{1}{e^{1/\lambda}+2} \& \mu_1 > \mu_2 e^{1/\lambda}$, then $p^\star_{a_1}(\mu_b, \mu_g) = 1$, $p^\star_{a_2}(\mu_b, \mu_g) = p^\star_{a_3}(\mu_b, \mu_g) = 0$*

The first point of Proposition E.1 states that if the lowest prior belief μ_3 is higher than a certain threshold, then the DM chooses among all three actions at optimality. Else, the DM chooses between two actions if μ_1 and μ_2 take sufficiently close values (second point of the proposition). Otherwise, the DM only chooses one state (third point). In this sense, Proposition E.1 characterizes the possible consideration sets of alternatives from which the DM chooses, as a function of priors μ_g and μ_b.

Figure E.1 illustrates these sets for two different information cost values. For instance, when both μ_g and μ_b are sufficiently low, the DM chooses $a = o$ without eliciting any signal. The regions labeled as "Only a" correspond to situations where the DM directly chooses action a and disregards the two others a priori. In these cases, the DM does not process any information. Similarly, the DM rules out the moderate state a priori and chooses $a = y$ and $a = n$ with positive probability (i.e., forms a consideration set consisting only the good and bad states) if her belief toward the moderate state is sufficiently weak (see the region labeled as "Both y&n"). Only in the region labeled as "All" does the DM consider all possible states at the same time. As indicated in Figure E.1, this region shrinks when λ increases as the DM becomes less willing to process information for weaker prior beliefs.

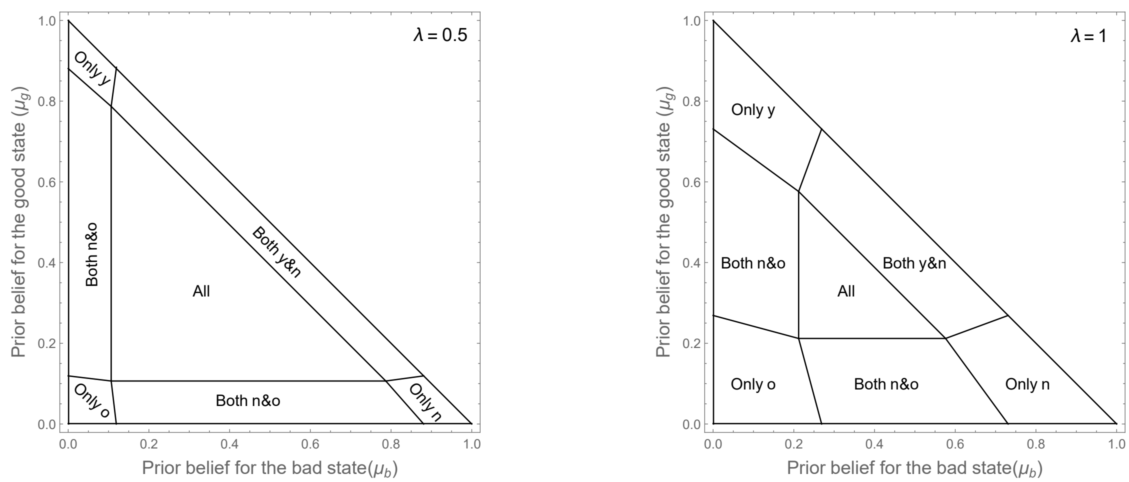


Figure E.1 Consideration set formation in $\mu_b \times \mu_g$ space

Using Proposition E.1 and following an approach similar to our base model, we can characterize the DM's decision accuracy, error probabilities and cognitive effort. These are detailed at the end of this section (see Appendix §E.1). In the following, we focus on how the presence of the machine impacts these metrics.

Upon a positive machine input, the DM rules out the bad state and updates her prior as $\mu_b^+ = 0$ and $\mu_g^+ = \frac{\mu_g}{P(X_1=+)}$. Similarly, upon a negative machine input, the DM rules out the good state with posterior beliefs $\mu_b^- = \frac{\mu_b}{P(X_1=-)}$ and $\mu_g^- = 0$. That is, the DM has always two states to consider in the presence of the machine: the moderate state and either the good or the bad state depending on the machine's input. Nonetheless, the machine always improves the DM's decision accuracy and expected value as we show in the next proposition.

Proposition E.2. *For any $\lambda > 0$, we have $A_m^\star \geq A^\star$ and $V_m^\star \geq V^\star$.*

Although the machine always increases overall accuracy, the machine may increase certain error types and induce the DM to exert more cognitive effort as in our base model. To explore this, we can follow our previous approach, and compare these different metrics with and without machine prediction using our characterizations in Appendix E.1. In particular, this consists in studying different scenarios, which depend on the different prior and posterior beliefs. However, the number of possible scenarios to consider grows exponentially with the number of possible states (from six possible cases in the base model with two states to 63 with three states.)[1] Although technically doable, exploring all these situations is tedious and of limited research interest. Instead, we focus on the situation in which the machine reduces uncertainty in a symmetric manner, i.e. where the DM's prior and posterior beliefs are symmetric. Specifically, we take $\mu_g = \mu_b = \mu < 0.5$ and $\pi(+,-) = \pi(-,+) = 1/2 - \mu$, so that the DM's posterior beliefs become $\mu_g^+ = \mu_b^- = 2\mu$ and the likelihood

[1] The inclusion of a new state gives rise to 7 different cases as depicted in Figure E.1 when the human DM is alone (as opposed to 3 in our base model which are "Only y", "Only n" and "Both y&n"). In the presence of the machine, we have 9 cases -3 cases for each of the two possible posterior beliefs given the machine's prediction on $X1$ (this is since the machine always eliminates one of the states for the DM). This requires studying up to 7x9=63 cases for a full-fledged analysis of the problem. By contrast, there are 3x3=9 cases in our base model - as there are only 3 cases to consider with the machine since $\mu^- = 0$. Some of these cases can further be ruled out since $\mu^+ > \mu$, which yields 6 possible cases for our base model as illustrated in Theorems 2, 3 and 4.

of a machine outcome is given by $P(X_1 = +) = P(X_1 = -) = 0.5$. We next characterize the impact of the machine on the DM's decision errors and cognitive effort in this set-up.

PROPOSITION E.3. *For any given* $\lambda > 0$, $\alpha_m^\star > \alpha^\star$ *if* $\frac{1}{2}\frac{1}{e^{1/\lambda}+1} < \mu < \mu_{fp}^*$ *and* $\alpha_m^\star \leq \alpha^\star$ *otherwise. Furthermore,* $\mu_{fp}^* < 1/3$, $\alpha_m^\star = \beta_m^\star$ *and* $\alpha^\star = \beta^\star$

Note that the false positive and negative errors are always equal, whether the machine is present or not ($\alpha_m^\star = \beta_m^\star$ and $\alpha^\star = \beta^\star$). This is because the DM's prior and posterior beliefs are symmetric in this set-up. In essence, Proposition E.3 states that the machine increases the DM's false positive (and negative) errors if she sufficiently favors the moderate state ($\mu_m > 1 - 2\mu_{fp}^* > 1/3$) a priori. This increase in the decision errors is actually offset by a decrease in *false moderate* errors, so that the machine always improves overall accuracy per Proposition E.2.

PROPOSITION E.4. *For any given* $\lambda > 0$, $C_m^\star > C^\star$ *if* $\frac{1}{2}\frac{1}{e^{1/\lambda}+1} < \mu < \mu_e^*$ *and* $C_m^\star \leq C^\star$ *otherwise. Furthermore,* $\mu_e^* < 1/3$.

Proposition E.4 shows that the machine increases the cognitive effort the DM exerts when she sufficiently favors the moderate state. In this case, the task difficulty increases with any machine input. To see this, recall that the DM without the machine tries to distinguish three states with initial belief $\mu_b = \mu_g = \mu$ and $\mu_m = 1 - 2\mu$. When the machine gives a positive (resp. negative) signal, the DM deals with two states with posterior $\mu_g = 2\mu$ (resp. $\mu_b = 2\mu$) and $\mu_m = 1 - 2\mu$. When μ is sufficiently small, distinguishing the remaining two states becomes more difficult for the DM.

Overall, Proposition E.3 and Proposition E.4 establish that our main results for the base model are robust to decision settings where the machine reduces uncertainty in a symmetric manner.

E.1. Accuracy and Cognitive Effort in the Presence of Three States

Accuracy and Decision Errors We denote by $p_{a|\omega}^\star(\mu_b, \mu_g)$ the posterior probability that the DM chooses action a given state ω. Accuracy is defined as the sum of joint probability of choosing $a = y$ in the good state, choosing $a = n$ in the bad state and choosing $a = o$ in the moderate state. These joint probabilities can be written in terms of the DM's optimal choice probabilities and corresponding posteriors. From Theorem 1 in Caplin et al. (2019), the DM's optimal posteriors are

$$p_{\omega_j|a_i}^\star = \begin{cases} \frac{e^{1/\lambda}}{e^{1/\lambda}+2} & \text{for } i = j \\ \frac{1}{e^{1/\lambda}+2} & \text{otherwise} \end{cases} \text{if } \mu_3 > \frac{1}{e^{1/\lambda}+2}$$

$$p_{\omega_j|a_i}^\star = \begin{cases} \frac{e^{1/\lambda}(\mu_1+\mu_2)}{e^{1/\lambda}+1} & \text{for } i = j \text{ and } j \leq 2 \\ \frac{(\mu_1+\mu_2)}{e^{1/\lambda}+1} & \text{for } i \neq j \text{ and } j \leq 2 \\ \mu_j & j = 3 \end{cases} \text{if } \mu_3 < \frac{1}{e^{1/\lambda}+2} \text{ and } \mu_1 < \mu_2 e^{1/\lambda}$$

$$p_{\omega_j|a_i}^\star = \mu_j \text{ if } \mu_3 < \frac{1}{e^{1/\lambda}+2} \text{ and } \mu_1 > \mu_2 e^{1/\lambda}$$

Plugging in these and after some algebra, we obtain the accuracy function as

$$\begin{aligned} A^\star(\mu_b, \mu_g) &= p_{g|y}^\star p_y^\star(\mu_b, \mu_g) + p_{b|n}^\star p_n^\star(\mu_b, \mu_g) + p_{m|o}^\star p_o^\star(\mu_b, \mu_g) \\ &= \begin{cases} \mu_1 & \text{if } \mu_3 < \frac{1}{e^{1/\lambda}+2} \text{ and } \mu_1 > \mu_2 e^{1/\lambda} \\ \frac{e^{1/\lambda}(\mu_1+\mu_2)}{e^{1/\lambda}+1} & \text{if } \mu_3 < \frac{1}{e^{1/\lambda}+2} \text{ and } \mu_1 < \mu_2 e^{1/\lambda} \\ \frac{e^{1/\lambda}}{e^{1/\lambda}+2} & \text{if } \mu_3 > \frac{1}{e^{1/\lambda}+2} \end{cases} \end{aligned} \tag{E.1}$$

The resulting characterization has many similarities with our base model with two actions and two states. Note that when only one option is selected, then DM's accuracy is just her prior for that state which, in fact, will be the one with the strongest prior belief. When all options are selected, expected accuracy does not depend on prior belief. Similar to our base model, this is related to the symmetry of DM's payoffs so that her posterior beliefs are also symmetric. On the other hand, when the DM selects only two options in optimality, expected accuracy is just the scaled-down version of the accuracy in our base model with the likelihood of both selected states (i.e., $(\mu_1+\mu_2)$). Note that $e^{1/\lambda}/\left(e^{1/\lambda}+1\right)$ is the expected accuracy in our base model when both options are selected. Due to this scaling, expected accuracy in this case depends on the DM's prior belief.

In a similar manner, the error probabilities (respectively false positive, false negative and false moderate) can be computed by plugging in choice probabilities and optimal posteriors as follows:

$$\alpha^\star(\mu_b,\mu_g) = p_y^\star(\mu_b,\mu_g)\left(1-p_{g|y}^\star\right)$$
$$\beta^\star(\mu_b,\mu_g) = p_n^\star(\mu_b,\mu_g)\left(1-p_{g|n}^\star\right)$$
$$\gamma^\star(\mu_b,\mu_g) = p_o^\star(\mu_b,\mu_g)\left(1-p_{g|o}^\star\right)$$

Cognitive Effort When there are three states, the entropy as a function of μ_g and μ_b is

$$H(\mu_b,\mu_g) = -\mu_g\log\mu_g - \mu_b\log\mu_b - (1-\mu_g-\mu_b)\log(1-\mu_g-\mu_b).$$

The DM's cognitive effort is then defined as

$$C^\star(\mu_b,\mu_g) = \lambda\left[H(\mu_b,\mu_g) - p_y^\star(\mu_b,\mu_g)H\left(p_{g|y}^\star,p_{b|y}^\star\right) - p_y^\star(\mu_b,\mu_g)H\left(p_{g|n}^\star,p_{b|n}^\star\right) - p_y^\star(\mu_b,\mu_g)H\left(p_{g|o}^\star,p_{b|o}^\star\right)\right]$$

which reduces (after some algebra) to

$$C^\star(\mu_b,\mu_g) = \begin{cases} 0 & \text{if } \mu_3 < \frac{1}{e^{1/\lambda}+2} \text{ and } \mu_1 > \mu_2 e^{1/\lambda} \\ \lambda\left[H(\mu_b,\mu_g) - (\mu_1+\mu_2)\varphi(\lambda) - H(\mu_1+\mu_2,0)\right] & \text{if } \mu_3 < \frac{1}{e^{1/\lambda}+2} \text{ and } \mu_1 < \mu_2 e^{1/\lambda} \\ \lambda\left[H(\mu_g,\mu_b) - \varphi^{(2)}(\lambda)\right] & \text{if } \mu_3 \geq \frac{1}{e^{1/\lambda}+2} \end{cases} \tag{E.2}$$

where

$$\varphi(\lambda) = \ln\left(e^{1/\lambda}+1\right) - \frac{1}{\lambda}\frac{e^{1/\lambda}}{e^{1/\lambda}+1}, \tag{E.3}$$

and

$$\varphi^{(2)}(\lambda) = -\frac{1}{\lambda}\frac{e^{\frac{1}{\lambda}}}{e^{\frac{1}{\lambda}}+2} + \ln\left(e^{\frac{1}{\lambda}}+2\right). \tag{E.4}$$

When the DM's prior is sufficiently strong toward a particular state, she does not process information and chooses the corresponding action. Then her cognitive effort is 0. When the DM chooses to process information about all states then her expected cognitive effort is in a very similar structure as our base model with two states. In particular, it is directly proportional to the difference between DM's prior entropy and expected posterior entropy given by $\varphi^{(2)}(\lambda)$. Note that it is independent of DM's prior belief and increasing in information cost λ. Lastly, when DM chooses to select two options and disregards one, then her expected posterior entropy depends on her prior belief.

Appendix F: The Impact of Restricting Information Sources for the DM

An information acquisition model for human decision maker which predicts that a free and perfectly accurate information (X_1 in our model) deteriorates her overall accuracy is not credible, at least in our view. This is, however, what a classical sequential hypothesis testing model as in DeGroot (1962) would predict in our context. In this section we illustrate this point. The key difference that the sequential testing introduces is that it limits the choice of signals (i.e. tests) that the DM can elicit.

Assume that at each period the DM may decide to conduct an imperfect test which provides a positive or negative signal about the true state. Each test costs K to the DM. There are infinite number of tests, but we assume that the DM decides optimally when to stop the search and commit to a decision. This corresponds to a standard optimal-stopping problem. To formulate this, we denote the machine signal at period k as $Z_k \in (+,-)$. We assume that the accuracy of the test is exogenous and known to the DM a priori. Assume that the test gives a signal with a true positive rate $\phi = P(Z_k = +|\omega = g)$ and false positive rate $\psi = P(Z_k = +|\omega = b)$. We denote the DM's prior belief that the state is good at period k as $\mu_k = P(\omega = g)$. If the DM decides to conduct the test, she updates her prior belief as following the Bayes' rule:

$$\mu_{k+1} = \begin{cases} \frac{\alpha\mu_k}{\alpha\mu_k+\beta(1-\mu_k)} & \text{if } Z_k = + \\ \frac{(1-\alpha)\mu_k}{(1-\alpha)\mu_k+(1-\beta)(1-\mu_k)} & \text{if } Z_k = - \end{cases}.$$

As in our base model, the DM aims to maximize her accuracy net of her total search cost. This is an optimal-stopping problem. We can write the corresponding DP formulation as

$$\begin{aligned} V_k(\mu_k) &= \max\{E_{Z_k}[V_{k+1}(\mu_{k+1})] - K, \max\{\mu_k, 1-\mu_k\}\} \\ V_T(\mu_T) &= \max\{\mu_T, 1-\mu_T\} \end{aligned}$$

where T is the terminal period. Here $E_{Z_k}[V_{k+1}(\mu_{k+1})] - \lambda$ is the DM's expected utility if she chooses to continue sampling where $P(Z_k = 1) = \alpha\mu_k + \beta(1-\mu_k)$. If she decides to choose an action and stop observing the signal, she obtains $\max\{\mu_k, 1-\mu_k\}$ by deciding based on her prior μ_k.

Let us denote DM's optimal information acquisition decision at period k as u_k^*, which can be written as

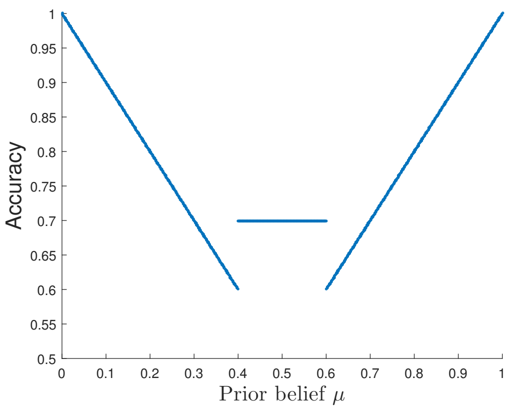
$$u_k^* = \begin{cases} 1 & \text{if } \max\{\mu_k, 1-\mu_k\} < E_Z[V_{k+1}(\mu_{k+1})] - K. \\ 0 & \text{otherwise} \end{cases}$$

where 1 is continue sampling and 0 is stop sampling. Accordingly, we can write the DM's decision accuracy function as

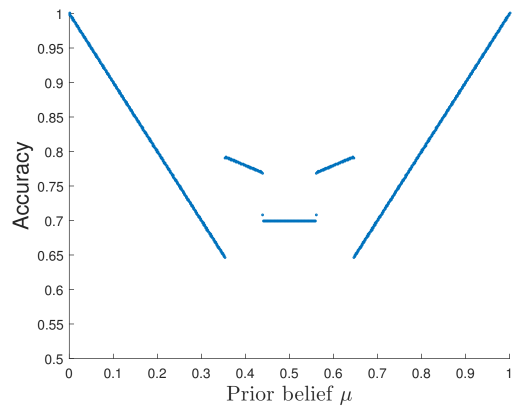
$$\begin{aligned} A_k(\mu_k) &= \begin{cases} P(Z_k = 1)\, A_{k+1}\left(\frac{\alpha\mu_k}{\alpha\mu_k+\beta(1-\mu_k)}\right) + P(Z_k = 0)\, A_{k+1}\left(\frac{(1-\alpha)\mu_k}{(1-\alpha)\mu_k+(1-\beta)(1-\mu_k)}\right) & \text{if } u_k^* = 1 \\ \max\{\mu_k, 1-\mu_k\} & \text{if } u_k^* = 0 \end{cases} \\ A_T(\mu_k) &= \max\{\mu_T, 1-\mu_T\}. \end{aligned}$$

As in our base model, we assume that the machine can reveal X_1 at no cost. Depending on the DM's updated prior upon the machine information, she can conduct further tests as explained above. We compare the DM's overall accuracy in these two cases (with and without the machine).

As a numerical example take $\phi = 0.3$, $\psi = 0.7$ and $K = 0.1$. Numerically solving the corresponding optimal stopping problem yields the accuracy function depicted in Figure F.1a for the DM. In this case, she chooses not to conduct any test if her prior for the good state μ is less than 0.4 or higher than 0.6. When her prior belief is sufficiently close to 0.5 ($\mu \in [0.4, 0.6]$), she chooses to conduct a single test and decides accordingly which yields an accuracy level of 0.7. Note that the DM's accuracy is neither convex nor continuous in prior belief unlike our model that is based on the rational inattention framework (please see Figure 2a).



(a) $K = 0.1$



(b) $K = 0.08$

Figure F.1 DM's decision accuracy as a function of prior belief ($\alpha = 0.7$, $\beta = 0.3$)

To see how the machine decreases the DM's accuracy in this scenario, take, for instance, $\mu = 0.5$. This yields an accuracy level of $A = 0.7$ per Figure F.1a. Assume further that the DM's posterior belief is $\mu^+ = 0.6$ when the machine provides a positive signal. This yields an ex-post accuracy of 0.6. Since the probability of a positive signal by the machine is $P(X_1 = +) = 0.5/0.6 = 5/6$, the DM's expected accuracy with the machine is $A_m = 5/6 * 0.6 + 1/6 * 1 = 0.667$ which is less than 0.7, the DM's accuracy without the machine.

Figure F.1b plots DM's accuracy function for a slightly lower test cost $K = 0.08$. In this case, the DM chooses to conduct more than one test when she is sufficiently uncertain as information cost is lower. Similarly, we can also find parameters in this scenario where machine can strictly decrease the DM's overall expected accuracy.

These simple counter-examples show that if the DM is able to elicit costly information from a *restrictive* set of signal sources, it is possible that a free and accurate machine information may actually reduce her overall accuracy.

References

Caplin A, Dean M, Leahy J (2019) Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies* 86(3):1061–1094.

Cover TM, Thomas JA (2012) *Elements of information theory* (John Wiley & Sons, Hoboken, NJ).

DeGroot MH (1962) Uncertainty, information, and sequential experiments. *The Annals of Mathematical Statistics* 33(2):404–419.

Matějka F, McKay A (2015) Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review* 105(1):272–98.

Özer Ö, Zheng Y, Chen KY (2011) Trust in forecast information sharing. *Management Science* 57(6):1111–1137.