Online Appendix to Speed, Accuracy, and the Optimal Timing of Choices

Drew Fudenberg* Philipp Strack† Tomasz Strzalecki‡

June 14, 2018

1 Additional Definitions of Accuracy

Our Definition 1 can be equivalently expressed in terms of the monotone likelihood ratio property. Let P be a choice process and let f^i be the density of F^i with respect to F. Wlog suppose that a is the correct choice and b is not. Suppose that F^a is absolutely continuous w.r.t. F^b; we say that F^a and F^b have the *monotone likelihood ratio property*, denoted $F^a \succsim_{\text{MLRP}} F^b$, if the likelihood $f^a(t)/f^b(t)$ is an increasing function.

Though the above concepts can be useful for theoretical analysis, in empirical work time periods will need to be binned to get useful test statistics. For this reason we introduce two weaker concepts that are less sensitive to finite samples, as their oscillation is mitigated by conditioning on larger sets of the form $[0,t]$. First, let $Q^i(t) := \frac{P(\{i\}\times[0,t])}{F(t)}$ be the probability of choosing i conditional on stopping in the interval $[0,t]$. Second, we say that F^a first order stochastically dominates F^b, denoted $F^a \succsim_{\text{FOSD}} F^b$ if $F^a(t) \leq F^b(t)$ for all $t \in T$. Below, we summarize the relationships between these concepts.

*Department of Economics, MIT

†Department of Economics, University of California, Berkeley

‡Department of Economics, Harvard University

Lemma O.1.

1. Let P be a choice process and suppose that F^a is absolutely continuous w.r.t. F^b. Then P has increasing (decreasing/constant) accuracy if and only if $F^a \succsim_{\text{MLRP}} F^b$ ($F^a \precsim_{\text{MLRP}} F^b$ / $F^a = F^b$).

2. If P has increasing (decreasing/constant) accuracy, then $Q^a(t)$ is an increasing (decreasing, constant) function.

3. If $Q^a(t)$ is an increasing (decreasing, constant) function, then $F^a \succsim_{\text{FOSD}} F^b$ ($F^a \precsim_{\text{FOSD}} F^b$, $F^a = F^b$).

Proof of Lemma O.1

To prove part (1) note that by the definition of a conditional distribution (property (c) p. 343 of Dudley, 2002) we have $F^i(t) = \int_{[0,t]} \frac{p^i(s)}{P^i} dF(s)$, so the density of F^i with respect to F is $f^i(t) = \frac{p^i(t)}{P^i}$. Since F^a is absolutely continuous w.r.t. F^b, the ratio $\frac{f^a(t)}{f^b(t)}$ is well defined F-almost everywhere and equals $\frac{p^a(t)}{p^b(t)} \frac{P^a}{P^b}$. This expression is increasing (decreasing, constant) if and only if $p^a(t)$ is increasing (decreasing, constant).

To prove part (2), note that by the definition of a conditional distribution we have

$$Q^i(t) = \frac{P^i F^i(t)}{F(t)} = \frac{\int_{[0,t]} p^i(s) dF(s)}{F(t)}. \tag{1}$$

Thus, for $t < t'$ we have

$$\begin{aligned} Q^l(t) > Q^l(t') \ \text{ iff } & \frac{\int_{[0,t]} p^l(s) dF(s)}{F(t)} \geq \frac{\int_{[0,t]} p^l(s) dF(s) + \int_{(t,t']} p^l(s) dF(s)}{F(t) + [F(t') - F(t)]} \\ \text{iff } & \frac{\int_{[0,t]} p^l(s) dF(s)}{F(t)} \geq \frac{\int_{(t,t']} p^l(s) dF(s)}{F(t') - F(t)}, \end{aligned}$$

which is true if $p^l(\cdot)$ is a decreasing function since the LHS is the average of p^l on $[0,t]$ and the RHS is the average on $(t,t']$. However, the opposite implication may not hold, for example, consider $p^l(t) := (t - 2/3)^2$ and $F(t) = t$ for $t \in [0,1]$. Then $p^l(t)$ is not decreasing, but $Q^l(t)$ is.

To prove part (3), note that by (1) we have

$$F^l(t) > F^r(t) \ \text{ iff } \ \frac{Q^l(t)}{P^l} \geq \frac{Q^r(t)}{P^r} \ \text{ iff } \ Q^l(t) \geq P^l = \lim_{s\to\infty} Q^l(s),$$

where we used the fact that $Q^l(t) + Q^r(t) = 1$ and $P^l + P^r = 1$. Thus, if Q^l is a decreasing function, the RHS will hold. However, the opposite obviously doesn't have to hold. □

2 A more general model of attention

Given attention levels $\beta_t^l, \beta_t^r \geq 0$ the signal at time t is given by

$$\mathrm{d}Z_t^i = \left(\beta_t^i\right)^{\gamma/2} \theta^i \mathrm{d}t + \mathrm{d}B_t^i\,,$$

for some fixed $\gamma \geq 1$. We assume that the attention budget is fixed at every point in time, $\beta_t^l + \beta_t^r \leq 2$. The posterior variance of alternative i is given by $\left(\sigma_t^i\right)^2 = \frac{1}{\left(\sigma_0^i\right)^{-2} + \int_0^t (\beta_s^i)^\gamma \mathrm{d}s}$, so the posterior variance on the difference is

$$v_t = \left(\sigma_t^l\right)^2 + \left(\sigma_t^r\right)^2 = \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^l} + \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^r}\,,$$

where $y_t^i = \int_0^t \left(\beta_s^i\right)^\gamma \mathrm{d}s$ is a measure of the total attention the agent has spent on alternative $i \in \{l, r\}$.

We first consider the auxiliary problem of minimizing the posterior variance at some *fixed time* t. At each point in time s the agent optimally exhausts the total attention budget of two, $\beta_t^s + \beta_s^r = 2$. We divide the proof in two cases $\gamma \geq 1$ and $\gamma \leq 1$.

$\underline{\gamma \leq 1}$: We claim that the agent can minimize the posterior variance v_t by paying equal attention to the two signals $\beta_t^l = \beta_t^r = 1$. To see this, note that, $y_t^l + y_t^r$ is bounded by $2t$

$$\begin{aligned} y_t^l + y_t^r &= \int_0^t \left(\beta_s^l\right)^\gamma + \left(\beta_s^r\right)^\gamma \mathrm{d}s \leq \int_0^t \max_{\beta^l, \beta^r, \beta^l + \beta^r \leq 2} \left[\left(\beta^l\right)^\gamma + \left(\beta^r\right)^\gamma\right] \mathrm{d}s \\ &= \int_0^t \left[(1)^\gamma + (1)^\gamma\right] \mathrm{d}s = \int_0^t 2\mathrm{d}s = 2t\,. \end{aligned}$$

Hence, a relaxed version of the problem of minimizing posterior variance at time t is given by

$$\begin{aligned} &\min_{(y_t^l, y_t^r)} \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^l} + \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^r} \\ \text{s.t.} \qquad & y_t^l + y_t^r = 2\,t\,. \end{aligned}$$

As the objective function is concave in y_t^l and y_t^r it follows that the solution to the above problem satisfies $y_t^l = y_t^r = t$. As this is achievable by $\beta_t^l = \beta_t^r = 1$ this means that the policy which pays equal attention minimizes the posterior variance at time any time t *simultaneously*.

$\underline{\gamma \geq 1}$: We claim that the agent can without loss use only attention levels 0 or 2. To see this, note that, the agent can always achieve the total attention y_t^i by paying attention only to i for the length of time $y_t^i/2^\gamma$

$$y_t^i = \int_0^t \left(\beta_s^i\right)^\gamma \mathrm{d}s = \int_0^{y_t^i/2^\gamma} 2^\gamma \mathrm{d}s\,.$$

It thus suffices to show that $y_t^l/2^\gamma + y_t^r/2^\gamma \leq t$, as this implies that by paying full attention to signal l from time $[0, y_t^l/2^\gamma]$ and then full attention to signal r for time $[y_t^l/2^\gamma, y_t^l/2^\gamma + y_t^r/2^\gamma]$ the agent can replicate the vector of total attention (y_t^l, y_t^r) . We have that the time it takes the agent to replicate the attention vector (y_t^l, y_t^r) is less than t

$$\begin{aligned} y_t^l/2^\gamma + y_t^r/2^\gamma &= \frac{1}{2^\gamma} \int_0^t \left(\beta_s^l\right)^\gamma + \left(\beta_s^r\right)^\gamma \mathrm{d}s \leq \frac{1}{2^\gamma} \int_0^t \max_{\beta^l, \beta^r, \beta^l + \beta^r \leq 2} \left[\left(\beta^l\right)^\gamma + \left(\beta^r\right)^\gamma\right] \mathrm{d}s \\ &= \frac{1}{2^\gamma} \int_0^t 2^\gamma \mathrm{d}s = t\,. \end{aligned}$$

Hence, the problem of minimizing posterior variance at time t can be rewritten as

$$\begin{aligned} &\min_{(y_t^l, y_t^r)} \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^l} + \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^r} \\ \text{s.t.} \qquad & y_t^l + y_t^r = 2^\gamma\, t\,. \end{aligned}$$

As the objective function is concave in y_t^l and y_t^r it follows that the solution to the above problem satisfies $y_t^l = y_t^r = 2^{\gamma-1}\, t$. This means that any policy where the agent pays full attention to the left alternative $\beta_s^l = 2$ for half the time $(\frac{t}{2})$ and full attention to the right alternative $\beta_s^r = 2$ for half the time $(\frac{t}{2})$ minimizes the posterior variance at time t.

Next, consider the policy which divides time in intervals of equal length Δ and pays full attention to the left alternative in the even intervals and full attention to the right alternative in the odd intervals. By construction $y_t^l = y_t^r \to 2^{\gamma-1}\, t$ when the length of the intervals $\Delta \searrow 0$ for *every* t. As a consequence the limit policy minimizes the posterior variance at all points in time *simultaneously*, and the limit variance at each t is $\tilde{v}_t^\star = 2\frac{1}{\left(\sigma_0^i\right)^{-2} + 2^{\gamma-1}\, t}$. Note that if $\gamma = 1$ which is the case stated in the main text, then the same posterior variance is achieved by the policy which pays equal attention at every point in time $\beta_t^l = \beta_t^r = 1$ for all t.

So far we argued that the agent can *simultaneously* minimize the posterior variance at every point in time by equalizing the posterior variance on the two options. The last step is to show that mimimizng the posterior variance at each time is optimal. Fix an attention strategy β and denote by $\mathbb{E}^\beta[\cdot]$ the associated expectation operator, and by $\mathbb{E}^{\beta^\star}[\cdot]$ the expectation operator associated with the limiting case where the agent switches attention instantaneously between the two objects. The optimal stopping policy τ is a solution to

$$\sup_\tau \mathbb{E}^\beta \left[\max\{X_\tau^l, X_\tau^r\} - c\tau\right] = \sup_\tau \mathbb{E}^\beta \left[\max\{X_\tau^l - X_\tau^r, 0\} - c\tau\right] + X_0^r \tag{2}$$

By the Dambis, Dubins, Schwarz Theorem (see, e.g., Theorem 1.6, chapter V of Revuz and Yor, 1999) there exists a Brownian motion $(B_\nu)_{\nu\in[0,\sigma_0^2]}$ such that $X_t^l - X_t^r = B_{\sigma_0^2 - v_t}$; this a time change where the new scale is the posterior variance. Furthermore, we can define the stochastic process $\phi_\nu := \inf\{t\colon\ \sigma_0^2 - v_t \geq \nu\}$. By eq. (2) the value of the agent is given by

$$\sup_\tau \mathbb{E}^\beta \left[\max\{X^l_\tau - X^r_\tau, 0\} - c\tau\right] + X^r_0 = \sup_\nu \mathbb{E}\left[\max\{B_\nu, 0\} - c\,\phi_\nu\right] + X^r_0\,.$$

Recall that we denoted by $(\tilde{v}_t)_{t\geq 0}$ the limiting posterior variance process if the agent quickly switches attention constantly between the two signals. As the posterior variance v_t is greater than the posterior variance $\tilde{v}_t$ we have that $\phi_\nu \geq \tilde{\phi}_\nu := \inf\{t\colon\ \sigma_0^2 - \tilde{v}_t \geq \nu\}$. It follows from $\phi_r \geq \tilde{\phi}_r$ that the value when using the attention strategy β is smaller that the value achieved in the limit when the agent constantly switches attention between the two signals

$$\begin{aligned}\sup_\tau \mathbb{E}^\beta \left[\max\{X^l_\tau, X^r_\tau\} - c\tau\right] &= \sup_\nu \mathbb{E}\left[B_\nu - c\,\phi_\nu\right] + X^r_0 \leq \sup_\nu \mathbb{E}\left[\max\{B_\nu, 0\} - c\,\tilde{\phi}_\nu\right] + X^r_0 \\ &= \sup_\tau \mathbb{E}^{\beta^\star}\left[\max\{X^l_\tau, X^r_\tau\} - c\tau\right]\,. \quad \square\end{aligned}$$

3 Proofs omitted in the main text

3.1 Proof of Lemma 4

To prove this Lemma we need the following result.

Lemma O.2.

$$\begin{aligned}&V(0,x^l, x^r, c\lambda, \sigma_0, \alpha) = \\ &\quad = \lambda^{-1} \sup_{\tau'} \mathbb{E}\left[\max\left\{\lambda x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} \mathrm{d}M^l_s, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} \mathrm{d}M^r_s\right\} - c\tau'\right],\end{aligned}$$

where M^i_t are Brownian motions.

Proof of Lemma O.2. By definition, $V(0, x^l, x^r, c\lambda, \sigma_0, \alpha)$ equals

$$\sup_\tau \mathbb{E}\left[\max\left\{x^l + \int_0^\tau \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} \mathrm{d}W^l_s, x^r + \int_0^\tau \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} \mathrm{d}W^r_s\right\} - c\lambda\,\tau\right]$$

and by simple algebra this equals

$$\lambda^{-1}\sup_{\tau}\mathbb{E}\left[\max\left\{\lambda x^l+\int_0^{\tau}\frac{\lambda\alpha^{-1}}{\sigma_0^{-2}+s\alpha^{-2}}\mathrm{d}W^l_{(s\lambda^2)\lambda^{-2}},\lambda x^r+\int_0^{\tau}\frac{\lambda\alpha^{-1}}{\sigma_0^{-2}+s\alpha^{-2}}\mathrm{d}W^r_{(s\lambda^2)\lambda^{-2}}\right\}-c\lambda^2\tau\right].$$

We now change the speed of time and apply Proposition 1.4 of Chapter V of Revuz and Yor (1999) with $C_s := s\lambda^{-2}$and $H_s := \frac{\alpha^{-1}\lambda}{\sigma_0^{-2}+\alpha^{-2}\lambda^{-2}s}$ (pathwise to the integrals with limits τ and $\tau\lambda^2$) to get

$$\lambda^{-1}\sup_{\tau}\mathbb{E}\left[\max\left\{\lambda x^l+\int_0^{\tau\lambda^2}\frac{\lambda\alpha^{-1}}{\sigma_0^{-2}+s\lambda^{-2}\alpha^{-2}}\mathrm{d}W^l_{s\lambda^{-2}},\lambda x^r+\int_0^{\tau\lambda^2}\frac{\lambda\alpha^{-1}}{\sigma_0^{-2}+s\lambda^{-2}\alpha^{-2}}\mathrm{d}W^r_{s\lambda^{-2}}\right\}-c\lambda^2\tau\right].$$

In the next step we apply the time rescaling argument to conclude that $M^i_r := \lambda W^i_{r\lambda^{-2}}$ is a Brownian motion, and $\tau' = \tau\lambda^2$ is a stopping time measurable in the natural filtration generated by M. This yields

$$\lambda^{-1}\sup_{\tau'}\mathbb{E}\left[\max\left\{\lambda x^l+\int_0^{\tau'}\frac{\alpha^{-1}}{\sigma_0^{-2}+s\lambda^{-2}\alpha^{-2}}\mathrm{d}M^l_s,\lambda x^r+\int_0^{\tau'}\frac{\alpha^{-1}}{\sigma_0^{-2}+s\lambda^{-2}\alpha^{-2}}\mathrm{d}M^r_s\right\}-c\tau'\right]$$

□

Proof of (13)

By Lemma 2, part 7, $V(t,x^l,x^r,c,\sigma_0,\alpha)=V(0,x^l,x^r,c,\sigma_t,\alpha)$, so

$$\begin{aligned}k^*(t,c,\sigma_0,\alpha)&=\inf\{x>0\colon 0=V(t,0,-x,c,\sigma_0,\alpha)\}\\&=\inf\{x>0\colon 0=V(0,0,-x,c,\sigma_t,\alpha)\}=k(0,c,\sigma_t,\alpha).\end{aligned}$$

Proof of (14)

By Lemma O.2, $V(0, x^l, x^r, c\lambda, \sigma_0, \alpha)$ equals

$$\lambda^{-1} \sup_{\tau'} \mathbb{E}\left[\max\left\{\lambda x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} \mathrm{d}M_s^l, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} \mathrm{d}M_s^r\right\} - c\tau'\right]$$

$$\lambda^{-1}\lambda^2 \sup_{\tau'} \mathbb{E}\left[\lambda^{-2} \max\left\{\lambda x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} \mathrm{d}M_s^l, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} \mathrm{d}M_s^r\right\} - c\tau'\lambda^{-2}\right]$$

$$= \lambda \sup_{\tau'} \mathbb{E}\left[\max\left\{\lambda^{-1} x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\lambda^2\sigma_0^{-2} + s\alpha^{-2}} \mathrm{d}M_s^l, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\lambda^2\sigma_0^{-2} + s\alpha^{-2}} \mathrm{d}M_s^r\right\} - c\tau'\lambda^{-2}\right]$$

$$= \lambda V(0, x_1\lambda^{-1}, x_2\lambda^{-1}, c\lambda^{-2}, \sigma_0/\lambda, \alpha).$$

By setting $\hat{c} = c\lambda$ we have $V(0, x^l, x^r, \hat{c}, \sigma_0, \alpha) = \lambda V(0, x_1\lambda^{-1}, x_2\lambda^{-1}, \hat{c}\lambda^{-3}, \sigma_0/\lambda, \alpha)$, so

$$\begin{aligned} k^*(0, c, \sigma_0, \alpha) &= \inf\{x > 0 \colon 0 = V(0, 0, -x, c, \sigma_0, \alpha)\} \\ &= \inf\{x > 0 \colon 0 = V(0, 0, -x\,\lambda^{-1}, c\,\lambda^{-3}, \sigma_0/\lambda, \alpha)\} \\ &= \lambda \inf\{y > 0 \colon 0 = V(0, 0, -y, c\,\lambda^{-3}, \sigma_0/\lambda, \alpha)\} = \lambda\, k^*(0, c\,\lambda^{-3}, \sigma_0/\lambda, \alpha). \end{aligned}$$

Setting $\tilde{\sigma}_0 = \sigma_0/\lambda$ gives the result.

Proof of (15)

First, observe that $V(t, x_1, x_2, c, \sigma_0, \lambda\alpha)$ equals

$$= \sup_{\tau \geq t} \mathbb{E}\left[\max\left\{x_1 + \int_t^{\tau} \frac{\lambda^{-1}\alpha^{-1}}{\sigma_0^{-2} + \lambda^{-2}\alpha^{-2}s} \mathrm{d}W_s^1, x_2 + \int_t^{\tau} \frac{\lambda^{-1}\alpha^{-1}}{\sigma_0^{-2} + \lambda^{-2}\alpha^{-2}s} \mathrm{d}W_s^2\right\} - c(\tau - t)\right]$$

$$= \lambda \sup_{\tau \geq t} \mathbb{E}\left[\max\left\{\lambda^{-1}x_1 + \int_t^{\tau} \frac{\alpha^{-1}}{\lambda^2\sigma_0^{-2} + \alpha^{-2}s} \mathrm{d}W_s^1, \lambda^{-1}x_2 + \int_t^{\tau} \frac{\alpha^{-1}}{\lambda^2\sigma_0^{-2} + \alpha^{-2}s} \mathrm{d}W_s^2\right\} - (c\lambda^{-1})(\tau - t)\right]$$

$$= \lambda V(t, \lambda^{-1}x_1, \lambda^{-1}x_2, \lambda^{-1}c, \lambda^{-1}\sigma_0, \alpha)\,.$$

Thus,

$$\begin{aligned} k^*(t,c,\sigma_0,\lambda\alpha) &= \inf\{x>0\colon 0 = V(t,0,-x,c,\sigma_0,\lambda\alpha)\} \\ &= \inf\{x>0\colon 0 = V(t,0,-\lambda^{-1}x,\lambda^{-1}c,\lambda^{-1}\sigma_0,\alpha)\} \\ &= \lambda \inf\{y>0\colon 0 = V(t,0,-y,\lambda^{-1}c,\lambda^{-1}\sigma_0,\alpha)\} = \lambda\, k^*(t,\lambda^{-1}c,\lambda^{-1}\sigma_0,\alpha). \end{aligned}$$

Proof of (16)

By Lemma O.2, $V(0,x^l,x^r,c\lambda,\sigma_0,\alpha)$ equals

$$\lambda^{-1}\sup_{\tau'}\mathbb{E}\left[\max\left\{\lambda x^l + \int_0^{\tau'}\frac{\alpha^{-1}}{\sigma_0^{-2}+s\lambda^{-2}\alpha^{-2}}\mathrm{d}M_s^l, \lambda x^r + \int_0^{\tau'}\frac{\alpha^{-1}}{\sigma_0^{-2}+s\lambda^{-2}\alpha^{-2}}\mathrm{d}M_s^r\right\} - c\tau'\right]$$

As observing a less noisy signal is always better, we have that for all $\lambda > 1$

$$\begin{aligned} V(0,x^l,x^r,c\lambda,\sigma_0,\alpha) &\geq \lambda^{-1}\sup_{\tau}\mathbb{E}\left[\max\left\{\lambda x^l + \int_0^{\tau}\frac{\alpha^{-1}}{\sigma_0^{-2}+s\alpha^{-2}}\mathrm{d}M_s^l, \lambda x^r + \int_0^{\tau}\frac{\alpha^{-1}}{\sigma_0^{-2}+s\alpha^{-2}}\mathrm{d}M_s^r\right\} - c\tau\right] \\ &= \lambda^{-1}V(0,\lambda x^l,\lambda x^r,c,\sigma_0,\alpha) \end{aligned}$$

This implies that $k^*(t,\lambda c,\sigma_0,\alpha) \geq \lambda^{-1}k^*(t,c,\sigma_0,\alpha)$ for all $\lambda > 1$

$$\begin{aligned} k^*(t,\lambda c,\sigma_0,\alpha) &= \inf\{x>0\colon 0 = V(t,0,-x,\lambda c,\sigma_0,\alpha)\} \\ &\geq \inf\{x>0\colon 0 = V(t,0,-x\lambda,c,\sigma_0,\alpha)\} \\ &= \lambda^{-1}\inf\{y>0\colon 0 = V(0,0,-y,c,\sigma_0,\alpha)\} = \lambda^{-1}k(t,c,\sigma_0,\alpha)\,. \end{aligned}$$

3.1.1 Additional Results

Let

$$\bar{k}(t,c,\sigma_0,\alpha) = \frac{1}{2c\alpha^2(\sigma_0^{-2}+\alpha^{-2}t)^2}.$$

Lemma O.3. $\bar{k}$ is the only function that satisfies (13)–(16) with equality, and $\bar{b}$, defined above Proposition 4, is the associated boundary in the signal space.

Proof of Lemma O.3: Notice that by equations (13), (15), (14), and (16), applied in that order, it follows that

$$\bar{k}(t,c,\sigma_0,\alpha) = \bar{k}(0,c,\sigma_t,\alpha) = \alpha\bar{k}(0,\alpha^{-1}c,\alpha^{-1}\sigma_t,1) = \sigma_t\bar{k}(0,\alpha^2 c\sigma_t^{-3},1,1)$$
$$= \alpha^{-2}c^{-1}\sigma_t^4\bar{k}(0,1,1,1) = \frac{\kappa}{c\alpha^2(\sigma_0^{-2}+\alpha^{-2}t)^2},$$

where $\kappa = \bar{k}(0,1,1,1)$. Since $\bar{b}(t,c,\sigma_0,\alpha) = \alpha^2\bar{k}(t,c,\sigma_0,\alpha)\sigma_t^{-2}$, it follows that $\bar{b}(t,c,\sigma_0,\alpha) = \frac{\kappa}{c(\sigma_0^{-2}+\alpha^{-2}t)}$. The fact that $\kappa = \frac{1}{2}$ follows from the proof of Proposition 4, as any other constant would result in a contradiction as $t \to \infty$. □

Remark 1. Proposition 4 says that $\bar{k}$ is a good approximation of k^* for large t. An intuition for why this is true is as follows: Inequality (16) becomes an equality if additional information does not have value, which is the case when the agent already learned a lot, which is the case when t is large. Thus, for large t, k^* almost satisfies (16) with equality, i.e., it is almost equal to $\bar{k}$.

3.2 Proof of Proposition 2

Let $\kappa := \mathbb{E}\big[\max\{\theta^l,\theta^r\}\big]$ and fix a stopping time τ. To show that

$$\mathbb{E}\left[-\mathbf{1}_{\{X_\tau^l \geq X_\tau^r\}}(\theta^r-\theta^l)^+ - \mathbf{1}_{\{X_\tau^r > X_\tau^l\}}(\theta^l-\theta^r)^+ - c\tau\right] = \mathbb{E}\left[\max\{X_\tau^l,X_\tau^r\} - c\tau\right] - \kappa,$$

the cost terms can be dropped. Let D be the difference between the expected payoff from the optimal decision and the expected payoff from choosing the correct action, $D := \mathbb{E}\big[\max\{X_\tau^l,X_\tau^r\}\big] - \mathbb{E}\big[\max\{\theta^l,\theta^r\}\big]$. By decomposing the expectation into two events,

$$D = \mathbb{E}\Big[\mathbf{1}_{\{X_\tau^l \leq X_\tau^r\}}(X_\tau^l - \max\{\theta^l,\theta^r\}) + \mathbf{1}_{\{X_\tau^r < X_\tau^l\}}(X_\tau^r - \max\{\theta^l,\theta^r\})\Big].$$

Plugging in the definition of X^i_τ and using the law of iterated expectations, this equals

$$\begin{aligned}
&\mathbb{E}\bigg[\mathbf{1}_{\{X^l_\tau \le X^r_\tau\}}(\mathbb{E}[\theta^l|\mathcal{F}_\tau] - \max\{\theta^l,\theta^r\}) + \mathbf{1}_{\{X^r_\tau < X^l_\tau\}}(\mathbb{E}[\theta^r|\mathcal{F}_\tau] - \max\{\theta^l,\theta^r\})\bigg] \\
&= \mathbb{E}\bigg[\mathbf{1}_{\{X^l_\tau \le X^r_\tau\}}(\mathbb{E}[\theta^l|\mathcal{F}_\tau] - \mathbb{E}[\max\{\theta^l,\theta^r\}|\mathcal{F}_\tau]) + \mathbf{1}_{\{X^r_\tau < X^l_\tau\}}(\mathbb{E}[\theta^r|\mathcal{F}_\tau] - \mathbb{E}[\max\{\theta^l,\theta^r\}|\mathcal{F}_\tau])\bigg] \\
&= \mathbb{E}\bigg[\mathbf{1}_{\{X^l_\tau \le X^r_\tau\}}\mathbb{E}[-(\theta^r-\theta^l)^+|\mathcal{F}_\tau] + \mathbf{1}_{\{X^r_\tau < X^l_\tau\}}\mathbb{E}[-(\theta^l-\theta^r)^+|\mathcal{F}_\tau]\bigg] \\
&= \mathbb{E}\bigg[-\mathbf{1}_{\{X^l_\tau \le X^r_\tau\}}(\theta^r-\theta^l)^+ - \mathbf{1}_{\{X^r_\tau < X^l_\tau\}}(\theta^l-\theta^r)^+\bigg].
\end{aligned}$$

3.3 Proof of Proposition 4

We rely on Bather's (1962) analysis of the Chernoff model, which by Proposition 2 applies to our model. Bather studies a model with zero prior precision. Since such an agent never stops instantaneously, all that matters is her beliefs at $t > 0$, which are well defined even in this case, and given by $\hat{X}^i_t = t^{-1}Z^i_t$ and $\hat{\sigma}_t^{-2} = t\alpha^{-2}$. In Section 6, p. 619 Bather (1962) shows that

$$k^*\big(t, c, \infty, \frac{1}{\sqrt{2}}\big)\sqrt{t} = \frac{1}{4\,c\,t^{3/2}} + O\left(\frac{1}{t^3}\right).$$

which implies that

$$k^*\big(t, c, \infty, \frac{1}{\sqrt{2}}\big) = \frac{1}{4\,c\,t^2} + O\left(\frac{1}{t^{7/2}}\right).$$

Fix $\alpha > 0$. By equation (15) we have $k^*(t, c, \infty, \alpha) = \alpha\sqrt{2}k^*(t, \frac{1}{\alpha\sqrt{2}}c, \infty, \frac{1}{\sqrt{2}})$. Thus,

$$k^*\big(t, c, \infty, \alpha\big) = \frac{1}{2\,c\alpha^{-2}\,t^2} + O\left(\frac{1}{t^{7/2}}\right).$$

This implies that there exists $T, \eta > 0$ such that for all $t > T$ we have

$$\left|k^*\big(t, c, \infty, \alpha\big) - \frac{1}{2\,c\alpha^{-2}\,t^2}\right| \le \frac{\eta}{t^{7/2}}.$$

Fix $\sigma_0 > 0$ and let $s := t - \alpha^2\sigma_0^{-2}$. This way, the agent who starts with zero prior precision and waits t seconds has the same posterior precision as the agent who starts with σ_0^2 and waits

s seconds.[1] Thus, by (13) we have $k^*\big(t, c, \infty, \alpha\big) = k^*\big(s, c, \sigma_0, \alpha\big)$, so

$$\left| k^*\big(s, c, \sigma_0, \alpha\big) - \frac{1}{2\, c\alpha^2\, (\alpha^{-2} s + \sigma_0^{-2})^2} \right| \leq \frac{\eta}{(s + \alpha^2 \sigma_0^{-2})^{7/2}}.$$

Finally, since $b^*(s, c, \sigma_0, \alpha) = \alpha^2 k^*(s, c, \sigma_0, \alpha)\sigma_s^{-2}$, we have

$$\left| b^*\big(s, c, \sigma_0, \alpha\big) - \frac{1}{2\, c\alpha^2\, (\alpha^{-2} s + \sigma_0^{-2})} \right| \leq \frac{\eta}{(s + \alpha^2 \sigma_0^{-2})^{5/2}}.$$

3.4 Proof of Proposition 5

By Lemma 4 we have that $k^*(0, c, \lambda\sigma_0, \alpha) = \lambda\, k^*(0, c\lambda^{-3}, \sigma_0, \alpha)$ for all $\lambda > 0$. Choosing $\lambda = c^{1/3}$ and $\sigma_0 = \tilde{\sigma}_0 c^{-1/3}$ yields $k^*(0, c, \tilde{\sigma}_0, \alpha) = c^{1/3}\, k^*(0, 1, \tilde{\sigma}_0 c^{-1/3}, \alpha)$. Thus for $c > 1$ by choosing $t = \alpha^2\, \tilde{\sigma}_0^{-2} \left(c^{2/3} - 1\right)$, we have $\tilde{\sigma}_t = \tilde{\sigma}_0 c^{-1/3}$ and using equation (13) yields

$$k^*(0, c, \tilde{\sigma}_0, \alpha) = c^{1/3}\, k^*(\alpha^2\, \tilde{\sigma}_0^{-2}(c^{2/3} - 1), 1, \tilde{\sigma}_0, \alpha)\,.$$

Consider the difference between the correct barrier k^* and the approximate barrier $\bar{k}$

$$\begin{aligned}\left|k^*(0, c, \tilde{\sigma}_0, \alpha) - \bar{k}(0, c, \tilde{\sigma}_0, \alpha)\right| &= c^{1/3} \left|k^*(\alpha^2\, \tilde{\sigma}_0^{-2} \left\{c^{2/3} - 1\right\}, 1, \tilde{\sigma}_0, \alpha) - \bar{k}(\alpha^2\, \tilde{\sigma}_0^{-2} \left\{c^{2/3} - 1\right\}, 1, \tilde{\sigma}_0, \alpha)\right| \\ &= \frac{\tilde{\sigma}_0^2}{\alpha^2} c^{1/3} \left|b^*(\alpha^2\, \tilde{\sigma}_0^{-2} \left\{c^{2/3} - 1\right\}, 1, \tilde{\sigma}_0, \alpha) - \bar{b}(\alpha^2\, \tilde{\sigma}_0^{-2} \left\{c^{2/3} - 1\right\}, 1, \tilde{\sigma}_0, \alpha)\right|\end{aligned}$$

Note that in the right-hand side above the barrier is only evaluated at cost equal to one. Hence, by Proposition 4 we have that there exists constants $\eta, \underline{c} > 0$ independent of c such that for all $c \geq \underline{c}$

$$\left|k^*(0, c, \tilde{\sigma}_0, \alpha) - \bar{k}(0, c, \tilde{\sigma}_0, \alpha)\right| \leq \frac{\tilde{\sigma}_0^2}{\alpha^2}\, c^{1/3}\, \frac{\eta}{(\tilde{\sigma}_0^{-2} + \tilde{\sigma}_0^{-2} \left\{c^{2/3} - 1\right\})^{5/2}} = \frac{\tilde{\sigma}_0^7}{\alpha^2}\, c^{1/3} \frac{\eta}{c^{5/3}} = \frac{\tilde{\sigma}_0^7}{\alpha^2}\, \frac{\eta}{c^{4/3}}\,.$$

Multiplying the left and the right sides by $\alpha^2 \sigma_t^{-2}$ gives $\left|b^*(0, c, \tilde{\sigma}_0, \alpha) - \frac{1}{2\,c}\right| \leq \tilde{\sigma}_0^5\, \frac{\eta}{c^{4/3}}$. □

[1] To see this, observe that $\sigma_s^2 = \frac{1}{\sigma_0^{-2} + \alpha^{-2} s^2} = \frac{1}{\alpha^{-2} t} = \hat{\sigma}_t^2$.

3.5 Proof of Theorem 6

Let $G = \{t_n\}_{n=1}^N$ be a finite set of times at which the agent is allowed to stop and denote by $\mathcal{T}$ all stopping times τ such that $\tau \in G$ almost surely. As we restrict the agent to stopping times in $\mathcal{T}$, the stopping problem becomes a discrete time optimal stopping problem. By Doob's optional sampling theorem we have that

$$\begin{aligned}\sup_\tau \mathbb{E}\left[\max\{X^l_\tau, X^r_\tau\} - d(\tau)\right] &= \sup_\tau \mathbb{E}\left[\frac{1}{2}\max\{X^l_\tau - X^r_\tau, X^r_\tau - X^l_\tau\} + \frac{1}{2}(X^l_\tau + X^r_\tau) - d(\tau)\right]\\ &= \sup_\tau \mathbb{E}\left[\frac{1}{2}|X^l_\tau - X^r_\tau| - d(\tau)\right] + \frac{1}{2}(X^l_0 + X^r_0),\end{aligned}$$

so any optimal stopping time also solves $\sup_\tau \mathbb{E}\left[|X^l_\tau - X^r_\tau| - 2\,d(\tau)\right]$. Define $\Delta_n = |X^l_{t_n} - X^r_{t_n}|$ for all $n = 1, \ldots, N$. Observe that $(\Delta_n)_{n=1,\ldots,N}$ is a one-dimensional discrete time Markov process. To prove that for every barrier k in the posterior mean space (X) there exists a cost function which generates k by Theorem 1 in Kruse and Strack (2015) it suffices to prove that:

1. there exists a constant C such that $\mathbb{E}[\Delta_{n+1}|\mathcal{F}_{t_n}] \leq C(1 + \Delta_n)$.

2. Δ_{n+1} is increasing in Δ_n in the sense of first order stochastic dominance.

3. $z(n, y) = \mathbb{E}[\Delta_{n+1} - \Delta_n \mid \Delta_n = y]$ is strictly decreasing in y.

Condition 1 keeps the value of continuing from exploding, which would be inconsistent with a finite boundary. Conditions 2 and 3 combined ensure that the optimal policy is a cut-off rule.

In both the certain- and uncertain-difference models the mapping between X and Z is one-to-one and onto, so this implies the desired result for b.

3.5.1 Certain-Difference DDM

Set $Z_t = Z^l_t - Z^r_t = (\theta'' - \theta')t + \sqrt{2}\alpha B_t$. Then

$$\begin{aligned}l_t &= \log\left(\frac{\mathbb{P}\left[\theta = \theta_l \,|\, \mathcal{F}_t\right]}{\mathbb{P}\left[\theta = \theta_r \,|\, \mathcal{F}_t\right]}\right) = \log\left(\frac{\mu}{1-\mu}\right) + \log\left(\frac{\exp(-(4\alpha^2 t)^{-1})(Z_t - (\theta'' - \theta')t)^2}{\exp(-(4\alpha^2 t)^{-1})(Z_t - (\theta' - \theta'')t)^2}\right)\\ &= \log\left(\frac{\mu}{1-\mu}\right) + \frac{Z_t(\theta'' - \theta')}{\alpha^2}.\end{aligned}$$

Denote by $p_n = \mathbb{P}[\theta = \theta_l \mid \mathcal{F}_{t_n}]$ the posterior probability that l is the better choice. The expected absolute difference of the two choices satisfies

$$\begin{aligned}\Delta_n &= |X^l_{t_n} - X^r_{t_n}| = |p_n(\theta'' - \theta') + (1-p_n)(\theta' - \theta'')| \\ &= |(2p_n - 1)(\theta'' - \theta')| = 2(\theta'' - \theta') \left| p_n - \frac{1}{2} \right| .\end{aligned}$$

Let $\psi_n := [Z^l_{t_n} - Z^r_{t_n}] - [Z^l_{t_{n-1}} - Z^r_{t_{n-1}}]$ denote the change in the signal from t_{n-1} to t_n. We have that the log likelihood is given by $l_{n+1} = l_n + \alpha^{-2}(\theta'' - \theta')\psi_{n+1}$. We thus have

$$\Delta_n = 2(\theta'' - \theta') \left| \frac{e^{l_n}}{1 + e^{l_n}} - \frac{1}{2} \right| = 2(\theta'' - \theta') \left(\frac{e^{|l_n|}}{1 + e^{|l_n|}} - \frac{1}{2} \right) . \tag{3}$$

(1): It is easily seen that $\mathbb{E}\left[\Delta_{n+1} | \mathcal{F}_{t_n}\right] \leq (\theta'' - \theta')$, so for C big enough, $\mathbb{E}[\Delta_{n+1}|\mathcal{F}_{t_n}] \leq C(1 + \Delta_n)$.
(2): To simplify notation we introduce $m_n = |l_n|$. The process $(m_n)_{n=1,\ldots,N}$ is Markov. More precisely, $m_{n+1} = |m_n + \alpha^{-2}(\theta'' - \theta')\psi_{n+1}|$ is folded normal with mean of the underlying normal distribution equal to

$$\begin{aligned} m_n + \alpha^{-2}(\theta'' - \theta')\mathbb{E}[\psi_{n+1}|l_n] &= |l_n| + \alpha^{-2}(\theta'' - \theta')\Delta_n(t_{n+1} - t_n) \\ &= m_n + \alpha^{-2}\Big(\frac{2e^{m_n}}{1 + e^{m_n}} - 1\Big)(\theta'' - \theta')^2(t_{n+1} - t_n) \end{aligned} \tag{4}$$

and variance

$$\begin{aligned} Var\left[m_n + \alpha^{-2}(\theta'' - \theta')\psi_{n+1}\right] &= \alpha^{-4}(\theta'' - \theta')^2 Var\left[\psi_{n+1}\right] \\ &= 2\alpha^{-4}(\theta'' - \theta')^2(t_{n+1} - t_n) . \end{aligned}$$

As argued in part (2) of the uncertain difference case, a folded normal random variable increases in the sense of first order stochastic dominance in the mean of the underlying normal distribution. As (4) increases in m_n it follows that m_{n+1} increases in m_n in the sense of first order stochastic dominance. By (3) $m_n = |l_n|$ is increasing in Δ_n and Δ_{n+1} is increasing in $m_{n+1} = |l_{n+1}|$ this completes the argument.
(3): It remains to show that $z(n, \Delta_n)$ is decreasing in Δ_n. As $(p_n)_{n=1,\ldots,M}$ is a martingale, and

moreover conditioning on p is equivalent to conditioning on $1-p$, we have that

$$\begin{aligned}
z(n,\Delta_n) &= \mathbb{E}\left[\Delta_{n+1}\,\middle|\,p_n = \frac{\Delta_n}{2(\theta''-\theta')}+\frac{1}{2}\right]-\Delta_n \\
&= 2(\theta''-\theta')\mathbb{E}\left[|p_{n+1}-\frac{1}{2}|\,\middle|\,p_n = \frac{\Delta_n}{2(\theta''-\theta')}+\frac{1}{2}\right]-\Delta_n \\
&= 2(\theta''-\theta')\mathbb{E}\left[p_{n+1}-\frac{1}{2}\,\middle|\,p_n = \frac{\Delta_n}{2(\theta''-\theta')}+\frac{1}{2}\right] \\
&\qquad + 2(\theta"-\theta')\mathbb{E}\left[2\max\left\{\frac{1}{2}-p_{n+1},0\right\}\,\middle|\,p_n = \frac{\Delta_n}{(\theta''-\theta')}+\frac{1}{2}\right]-\Delta_n\,.
\end{aligned}$$

As p is a martingale we can replace p_{n+1} by p_n

$$\begin{aligned}
z(n,\Delta_n) &= 2(\theta''-\theta')\mathbb{E}\left[p_n-\frac{1}{2}\,\middle|\,p_n = \frac{\Delta_n}{2(\theta''-\theta')}+\frac{1}{2}\right] \\
&\qquad + 2(\theta''-\theta')\mathbb{E}\left[2\max\left\{\frac{1}{2}-p_{n+1},0\right\}\,\middle|\,p_i = \frac{\Delta_n}{2(\theta''-\theta')}+\frac{1}{2}\right]-\Delta_n \\
&= \Delta_n + 2(\theta''-\theta'))\mathbb{E}\left[2\max\left\{\frac{1}{2}-p_{n+1},0\right\}\,\middle|\,p_n = \frac{\Delta_n}{2(\theta''-\theta')}+\frac{1}{2}\right]-\Delta_n \\
&= 4(\theta''-\theta')\mathbb{E}\left[\max\left\{\frac{1}{2}-p_{n+1},0\right\}\,\middle|\,p_n = \frac{\Delta_n}{2(\theta''-\theta')}+\frac{1}{2}\right]\,.
\end{aligned}$$

The above term is strictly decreasing in Δ_n as p_{n+1} increases in the sense of first order stochastic dominance in p_n and p_n in the conditional expectation is increasing in Δ_n.

3.5.2 Uncertain-Difference DDM

Let us further define $\beta_i^2 = 2\sigma_{t_i}^2 - 2\sigma_{t_{i+1}}^2$. As $X_{t_i+1}^l - X_{t_i+1}^r$ is Normal distributed with variance β_i^2 and mean Δ_i we have that Δ_{i+1} is folded normal distributed with mean

$$\mathbb{E}_i\left[\Delta_{i+1}\right] = \beta_i\sqrt{\frac{2}{\pi}}e^{-\frac{\Delta_i^2}{2\beta_i^2}} + \Delta_i(1-2\Phi(\frac{-\Delta_i}{\beta_i}))\,,$$

where Φ denotes the normal cdf. Thus, the expected change in delta is given by

$$z(i,y) = \beta_i\sqrt{\frac{2}{\pi}}e^{-\frac{y^2}{2\beta_i^2}} - 2\,y\,\Phi(\frac{-\Delta_i}{\beta_i}).$$

(1): It is easily seen that $\mathbb{E}_i\left[\Delta_{i+1}\right] \leq \beta_i\sqrt{\frac{2}{\pi}} + \Delta_i$.

(2): As Δ_i is folded normal distributed we have that

$$\mathbb{P}_i(\Delta_{i+1} \leq y) = \frac{1}{2}\left[\operatorname{erf}\left(\frac{y+\Delta_i}{\beta_i}\right) + \operatorname{erf}\left(\frac{y-\Delta_i}{\beta_i}\right)\right].$$

Taking derivatives gives that

$$\frac{\partial}{\partial \Delta_i}\mathbb{P}_i(\Delta_{i+1} \leq y) = \frac{1}{2}\left[e^{-\left(\frac{y+\Delta_i}{\beta_i}\right)^2} - e^{-\left(\frac{y-\Delta_i}{\beta_i}\right)^2}\right] = \frac{1}{2}e^{-\left(\frac{y-\Delta_i}{\beta_i}\right)^2}\left[e^{-\frac{4\Delta_i y}{\beta_i}} - 1\right] < 0\,.$$

As $\Delta_i = |X^l_{t_i} - X^r_{t_i}|$ it follows that $y \geq 0$ and hence, Δ_{i+1} is increasing in Δ_i in the sense of first order stochastic dominance.

(3): The derivative of the expected change of the process Δ equals

$$\frac{\partial}{\partial y}z(i,y) = \frac{\partial}{\partial y}\left(\beta_i\sqrt{\frac{2}{\pi}}e^{-\frac{y^2}{2\beta_i^2}} - 2\,y\,\Phi\left(\frac{-\Delta_i}{\beta_i}\right)\right) = -2\Phi\left(\frac{-y}{\beta_i}\right) < 0\,.$$

Hence, z is strictly decreasing in y. □

3.6 Proof of Proposition 6

Recall that the analyst observes a DDM P for some known value of δ. First, by equation (4) in the paper we have $\frac{p^l(t)}{p^r(t)} = \exp\left(\frac{\delta b(t)}{\alpha^2}\right)$. Thus, $\frac{b(t)}{\alpha^2}$ is identified by

$$\frac{b(t)}{\alpha^2} = \frac{1}{\delta}\log\left(\frac{p^l(t)}{p^r(t)}\right) \tag{†}$$

Second, by equation (2) in the paper we have $\frac{Z_\tau}{\alpha^2} = \operatorname{sgn}(Z_\tau)\frac{b(\tau)}{\alpha^2}$. By equation (1) in the paper, $\frac{Z_\tau}{\alpha^2} = \frac{\delta}{\alpha^2}\tau + \frac{\sqrt{2}}{\alpha}B_\tau$. Combining these two equations and taking expectations, it follows from Doob's optional sampling theorem that

$$\frac{\delta}{\alpha^2}\mathbb{E}\left[\tau\right] = \mathbb{E}\left[\operatorname{sgn}(Z_\tau)\frac{b(\tau)}{\alpha^2}\right] \tag{‡}$$

Combining (†) and (‡) yields

$$\alpha^2 = \frac{\delta \mathbb{E}\left[\tau\right]}{\mathbb{E}\left[\operatorname{sgn}(Z_\tau)\frac{1}{\delta}\log\left(\frac{p^l(\tau)}{p^r(\tau)}\right)\right]} = \frac{\delta^2 \mathbb{E}\left[\tau\right]}{\mathbb{E}\left[\operatorname{sgn}(Z_\tau)\log\left(\frac{p^l(\tau)}{p^r(\tau)}\right))\right]}.$$

Note that as the agent's decision is observable, the sign of Z_τ is observable, so the right hand side is uniquely pinned down from P. Thus we can identify α^2. This immediately implies that we can identify $b(t)$, as

$$b(t) = \frac{1}{\delta}\log\left(\frac{p^l(t)}{p^r(t)}\right)\frac{\delta^2 \mathbb{E}\left[\tau\right]}{\mathbb{E}\left[\operatorname{sgn}(Z_\tau)\log\left(\frac{p^l(\tau)}{p^r(\tau)}\right)\right]} = \frac{\delta\,\log\left(\frac{p^l(t)}{p^r(t)}\right)\,\mathbb{E}\left[\tau\right]}{\mathbb{E}\left[\operatorname{sgn}(Z_\tau)\log\left(\frac{p^l(\tau)}{p^r(\tau)}\right)\right]}.$$

Proof of clause 1. Let $b(t) = \tilde{b}(t; g, h) = \frac{1}{g+ht}$. Suppose that $\tilde{b}(t; g, h) = \tilde{b}(t; \hat{g}, \hat{h})$ for all t. Then by setting $t = 0$ we get $g = \hat{g}$ and by setting $t = 1$ we get $h = \hat{h}$.

Proof of clause 2. Let $b(t) = \check{b}(t; g, h) = g\exp(-ht)$. Suppose that $\check{b}(t; g, h) = \check{b}(t; \hat{g}, \hat{h})$ for all t. Then by setting $t = 0$ we get $g = \hat{g}$ and by setting $t = 1$ we get $h = \hat{h}$.

Proof of clause 3. Let $b(t) = b^*(t, c, \sigma_0, \alpha)$. Suppose that $b^*(t, c, \sigma_0, \alpha) = b^*(t, \hat{c}, \hat{\sigma}_0, \alpha)$ for all t. Proposition 4, there exist $\eta, \hat{\eta}, T, \hat{T}$ such that

$$|\bar{b}(t, c, \sigma_0, \alpha) - b^*(t, c, \sigma_0, \alpha)| \leq \frac{\eta}{(\sigma_0^{-2} + \alpha^{-2}t)^{5/2}}$$

for $t > T$ and

$$|\bar{b}(t, \hat{c}, \hat{\sigma}_0, \alpha) - b^*(t, \hat{c}, \hat{\sigma}_0, \alpha)| \leq \frac{\hat{\eta}}{(\hat{\sigma}_0^{-2} + \alpha^{-2}t)^{5/2}}$$

for $t > \hat{T}$. Thus, for $t > \bar{T} := \max\{T, \hat{T}\}$ both of these inequalities hold, so

$$|\bar{b}(t, c, \sigma_0, \alpha) - \bar{b}(t, \hat{c}, \hat{\sigma}_0, \alpha)| \leq \frac{\eta}{(\sigma_0^{-2} + \alpha^{-2}t)^{5/2}} + \frac{\hat{\eta}}{(\hat{\sigma}_0^{-2} + \alpha^{-2}t)^{5/2}}. \tag{*}$$

Wlog, suppose that $\sigma_0 \geq \hat{\sigma}_0$. This implies that

$$\frac{\eta}{(\sigma_0^{-2}+\alpha^{-2}t)^{5/2}} + \frac{\hat{\eta}}{(\hat{\sigma}_0^{-2}+\alpha^{-2}t)^{5/2}} \leq \frac{\eta+\hat{\eta}}{(\sigma_0^{-2}+\alpha^{-2}t)^{5/2}};$$

thus, (*) implies that

$$|\bar{b}(t,c,\sigma_0,\alpha) - \bar{b}(t,\hat{c},\hat{\sigma}_0,\alpha)| \leq \frac{\bar{\eta}}{(\sigma_0^{-2}+\alpha^{-2}t)^{5/2}} \qquad (**)$$

for all $t > \bar{T}$.

By plugging in the formulas for $\bar{b}$, (**) becomes

$$\left|\frac{1}{2c(\sigma_0^{-2}+\alpha^{-2}t)} - \frac{1}{2\hat{c}(\hat{\sigma}_0^{-2}+\alpha^{-2}t)}\right| \leq \frac{\bar{\eta}}{(\sigma_0^{-2}+\alpha^{-2}t)^{5/2}}$$

for all $t > \bar{T}$. Rearranging terms yields

$$\frac{|\hat{c}(\hat{\sigma}_0^{-2}+\alpha^{-2}t) - c(\sigma_0^{-2}+\alpha^{-2}t)|}{c\hat{c}(\sigma_0^{-2}+\alpha^{-2}t)(\hat{\sigma}_0^{-2}+\alpha^{-2}t)} \leq \frac{2\bar{\eta}}{(\sigma_0^{-2}+\alpha^{-2}t)^{5/2}}.$$

Further rearranging yields

$$\frac{1}{c\hat{c}}|(\hat{c}\hat{\sigma}_0^{-2} - c\sigma_0^{-2}) + \alpha^{-2}t(\hat{c}-c)| \leq \frac{2\bar{\eta}(\hat{\sigma}_0^{-2}+\alpha^{-2}t)}{(\sigma_0^{-2}+\alpha^{-2}t)^{3/2}}. \qquad (***)$$

Dividing both sides by t gives

$$\frac{1}{c\hat{c}}\left|\frac{1}{t}(\hat{c}\hat{\sigma}_0^{-2} - c\sigma_0^{-2}) + \alpha^{-2}(\hat{c}-c)\right| \leq \frac{2\bar{\eta}(\frac{\hat{\sigma}_0^{-2}}{t}+\alpha^{-2})}{(\sigma_0^{-2}+\alpha^{-2}t)^{3/2}}.$$

Taking the limit $t \to \infty$, the LHS is $\frac{\alpha^{-2}(\hat{c}-c)}{c\hat{c}}$ and the RHS is zero, which implies that $\hat{c} = c$.

Given that, (***) becomes

$$|\hat{\sigma}_0^{-2} - \sigma_0^{-2}| \leq \frac{2c\bar{\eta}(\hat{\sigma}_0^{-2}+\alpha^{-2}t)}{(\sigma_0^{-2}+\alpha^{-2}t)^{3/2}}. \qquad (***)$$

Taking the limit as $t \to \infty$ the RHS is zero (by de L'Hopital's rule), which implies that $\hat{\sigma}_0 = \sigma_0$.

4 Numerical Methods

This section describes the details of the computation of the likelihood function. This is broken into two steps. First, we compute the optimal boundary $k^*(t, c, \sigma_0^2, \alpha)$. Second, given this boundary we compute the decision probabilities (using equation (4) in the paper) and hitting times (by simulation).

4.1 Computation of the optimal boundary k^*

To avoid unnecessary computations, we note that we without loss of generality we can fix $c = \alpha = 1$ and only vary σ_0. The next lemma follows from Lemma 4.

Lemma O.4.

$$k(t, c, \sigma_0, \alpha) = \alpha^{2/3} c^{1/3} h[(\sigma_0^{-2} + \alpha^{-2} t)^{-1/2} \alpha^{-2/3} c^{-1/3}],$$

where $h(x) := k^*(0, 1, x, 1)$.

Proof. By Lemma 4 we have

$$\begin{aligned}
k^*(t, c, \sigma_0, \alpha) &= \alpha k^*(t, \alpha^{-1} c, \alpha^{-1} \sigma_0, 1) \\
&= \alpha k^*(0, \alpha^{-1} c, (\alpha^2 \sigma_0^{-2} + t)^{-1/2}, 1) \\
&= \alpha \alpha^{-1/3} c^{1/3} k^*(0, 1, \alpha^{1/3} c^{-1/3} (\alpha^2 \sigma_0^{-2} + t)^{-1/2}, 1) \\
&= \alpha^{2/3} c^{1/3} k^*(0, 1, \alpha^{-2/3} c^{-1/3} (\sigma_0^{-2} + \alpha^{-2} t)^{-1/2}, 1)
\end{aligned}$$

where the first equality follows from (13), the second one from (11), the third one from (12), and the fourth one rearranges the terms. □

Thus, we just need to compute the h function. To do this, we first use a time change as in the proof of Lemma 2, part 6.

Lemma O.5.

$$V(t, x, 0, 1, \sigma_0, 1) = \sup_{\tau \geq \psi(t)} \mathbb{E}\left[\max\{W_\tau, 0\} - (q(\tau) - q(\psi(t)) \mid W_{\psi(t)} = x\right], \tag{5}$$

where $\psi(t) = \frac{2\sigma_0^2 t}{\sigma_0^{-2}+t}$, $q(s) = \psi^{-1}(s) = \frac{s\sigma_0^{-2}}{2\sigma_0^2 - s}$ and $(W_s) = (X_{q(s)})_{s\in[0,2\sigma_0^2]}$ is a Brownian motion.

Now we consider an approximation of the original optimization problem of the agent. We discretize both time and the posterior mean values and rely on a binomial approximation w_s of W_s. That is, $w_{s+\Delta_s}$ takes values $w_s - \Delta_w$ and $w_s + \Delta_w$ with equal probabilities. The discrete time version of the optimization problem is:

$$v(s, w, 0, 1, \sigma_0, 1) = \max\left\{\max\{w, 0\}, \frac{1}{2}v(s + \Delta_s, w + \Delta_w) + \frac{1}{2}v(s + \Delta_s, w - \Delta_w) - q'(s)\Delta_s\right\} \tag{6}$$

The solution to (6) approaches the solution to (5) when $\Delta_s, \Delta_w \to 0$, as long as $\Delta_s < \Delta_w^2$.

Fix σ_0 and let $\check{k}(s, 1, \sigma_0, 1)$ for $s \in [0, 2\sigma_0^2]$ be the optimal boundary in the above transformed problem. Then by equation (12) from Lemma 4 we have

$$\begin{aligned} h(x) &= k^*(0, 1, x, 1) = k^*(x^{-2} - \sigma_0^{-2}, 1, \sigma_0, 1) = \check{k}(\psi(x^{-2} - \sigma_0^{-2}), 1, \sigma_0, 1) \\ &= \check{k}(2\sigma_0^2 - 2x^2, 1, \sigma_0, 1) \end{aligned}$$

The Matlab code `solver_new9.m` solves for $\check{k}(s, 1, \sigma_0, 1)$ and outputs the values of the h function as a vector of prescribed length.

4.1.1 Robustness Checks

We implement two checks: first we verify if our solution is below the upper bound $\bar{k}$. This is indeed the case; moreover for values of σ_0 lower than one, the two functions are close to each other, so $\bar{k}$ is a good approximation, see Figure 1.

Second, we verify that our solution is above the lower bound from Bather, given by his implicit equation (6.4). Equation (6.4) of Bather gives the lower bound ρ as a solution to an implicit equation

$$ct^{\frac{3}{2}} = \frac{(1 + \rho^2)\phi(\rho) - 2\pi^{-\frac{1}{2}} e^{-\frac{1}{2}\rho^2}}{\rho + \frac{\psi(\rho)}{\rho}}$$

where $\phi(s) = (2\pi)^{-\frac{1}{2}} \int_s^\infty e^{-\frac{1}{2}y^2} dy$ and $\psi(s) = se^{-\frac{1}{2}s^2} \int_0^s e^{\frac{1}{2}y^2} dy$.

Note that $\phi(s) = 1 - \Phi(s)$, where Φ is the cdf of the standard Normal. Note also that

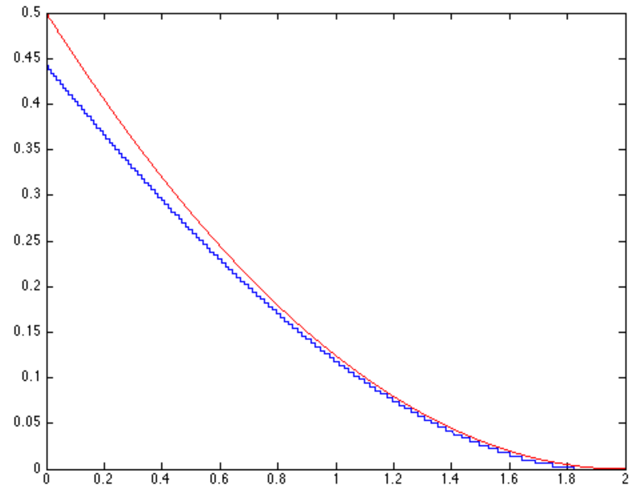


Figure 1: The function $\check{k}$ computed for $\sigma_0 = 1$ (blue) and the function $\bar{k}$ (red).

$\psi(s) = se^{-\frac{1}{2}s^2}\int_0^s e^{\frac{1}{2}y^2}dy = \sqrt{\frac{\pi}{2}}se^{-\frac{1}{2}s^2}\text{erfi}\left(\frac{s}{\sqrt{2}}\right)$, where erfi is the imaginary error function. Both of these functions are hard coded into Matlab, which makes things easier. Thus, we have

$$\frac{1}{t} = \left(\frac{(1+\rho^2)(1-\Phi(\rho)) - 2\pi^{-\frac{1}{2}}e^{-\frac{1}{2}\rho^2}}{c\left(\rho + \sqrt{\frac{\pi}{2}}e^{-\frac{1}{2}\rho^2}\text{erfi}\left(\frac{\rho}{\sqrt{2}}\right)\right)}\right)^{-\frac{2}{3}}.$$

Bather uses a different parametrization, so we need to express his equation in terms of our variables and solve it. We need to change variables from (t,ρ) to (s,x) in order to plot this function against our bound. The first change of variables is: $t = \frac{1}{\sigma_0^2 - 0.5s}$ and gives us

$$s = 2\sigma_0^2 - 2\left(\frac{(1+\rho^2)(1-\Phi(\rho)) - 2\pi^{-\frac{1}{2}}e^{-\frac{1}{2}\rho^2}}{c\left(\rho + \sqrt{\frac{\pi}{2}}e^{-\frac{1}{2}\rho^2}\text{erfi}\left(\frac{\rho}{\sqrt{2}}\right)\right)}\right)^{-\frac{2}{3}}.$$

The other change of variables we need to make is $\rho = x\sqrt{0.5t} = x\sqrt{\frac{1}{2\sigma_0^2 - s}}$. This gives us:

$$s - 2\sigma_0^2 + 2\left(\frac{(1+\frac{x^2}{2\sigma_0^2 - s})(1-\Phi(x\sqrt{\frac{1}{2\sigma_0^2-s}})) - 2\pi^{-\frac{1}{2}}e^{-\frac{x^2}{4\sigma_0^2-2s}}}{c\left(x\sqrt{\frac{1}{2\sigma_0^2-s}} + \sqrt{\frac{\pi}{2}}e^{-\frac{1}{2}\frac{x^2}{2\sigma_0^2-s}}\text{erfi}\left(x\sqrt{\frac{1}{4\sigma_0^2-2s}}\right)\right)}\right)^{-\frac{2}{3}} = 0.$$

Now, this relationship implicitly defines a function $s \mapsto x$, which we need to solve for numerically and plot against our bound, as in Figure 2.

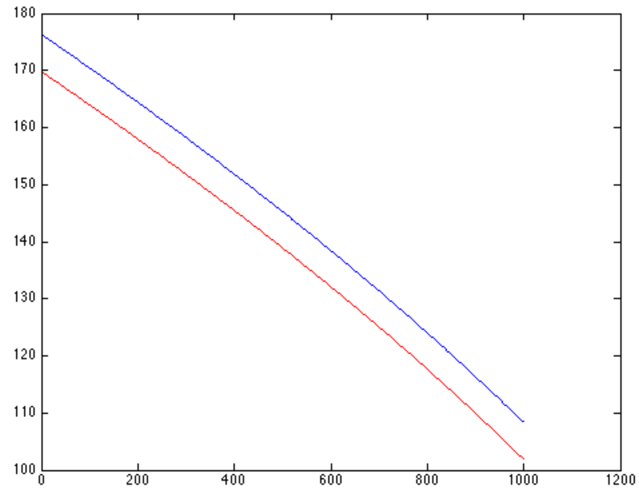


Figure 2: The function $\check{k}$ computed for $\sigma_0 = 30$ (blue) and the implicit function $s \mapsto x$ (red).

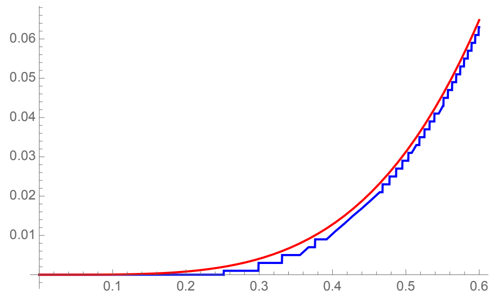
4.1.2 The function h used to compute the likelihood function

For any σ_0 let $h_{\sigma_0}(x) := \hat{k}(2\sigma_0^2 - 2x^2, 1, \sigma_0, 1)$, where $\hat{k}$ is the boundary computed by invoking the code solver.m with the value of σ_0. Moreover, define the function $\bar{h}(x) := \bar{k}(0, 1, x, 1) = 0.5x^4$. We compute the function h as follows:

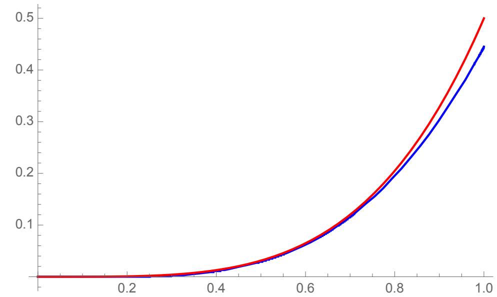
$$h(x) = \begin{cases} \bar{h}(x) & \text{if } x \in [0, .6) \\ h_1(x) & \text{if } x \in [.6, 1) \\ h_2(x) & \text{if } x \in [1, 2) \end{cases}$$

The following graphs illustrate the various component functions and provide the justification for choosing the particular cutoffs used in the above definition.

Note that $\bar{h}$ lies pointwise above h_1, which is consistent with $\bar{k}$ being an upper bound for k^*. Moreover, $\bar{h}$ is a very good approximation for h_1 for $x \in [0, .6)$ whereas for values of x above 0.6 the approximation is less good, which prompts us to rely on the numerically computed h_1.

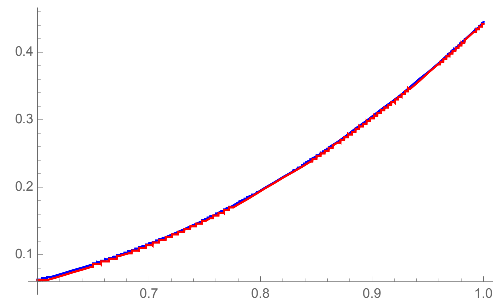


(a) $\bar{h}$ (red) and h_1 (blue) for $x \in [0, .6)$

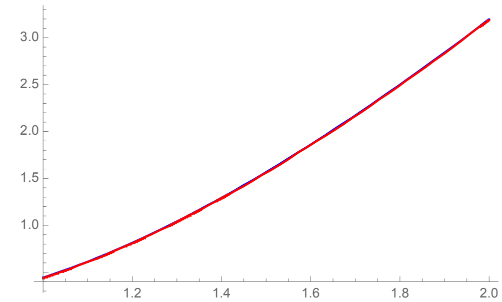


(b) $\bar{h}$ (red) and h_1 (blue) for $x \in [0, 1)$

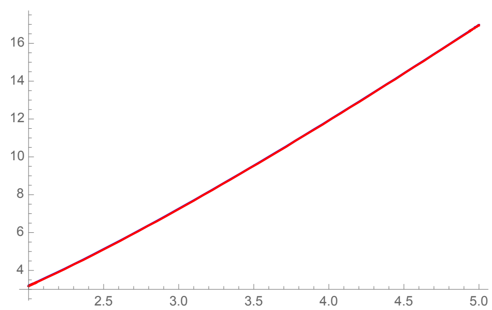
The following graph shows that the consecutive h_{σ_0} functions lie pretty much on top of each other.



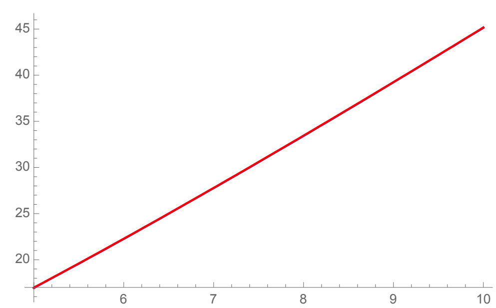
(c) h_1 (red) and h_2 (blue) for $x \in [.6, 1)$



(d) h_2 (red) and h_5 (blue) for $x \in [1, 2)$

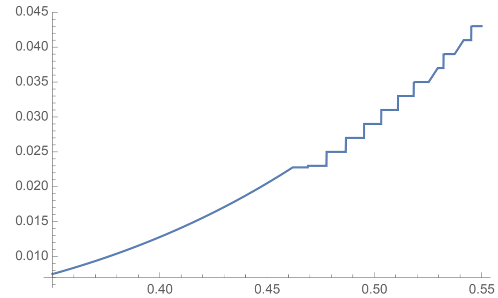


(e) h_5 (red) and h_{10} (blue) for $x \in [2, 5)$

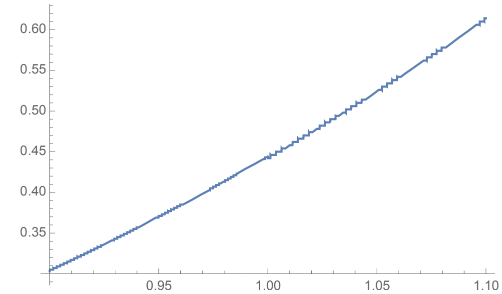


(f) h_{10} (red) and h_{30} (blue) for $x \in [5, 10)$

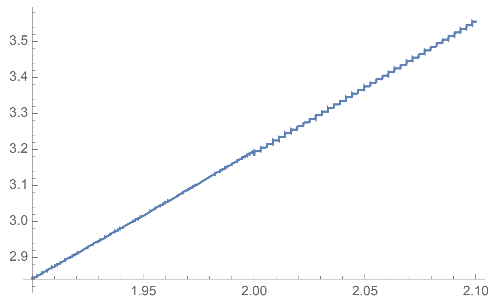
The following graph displays the h function around the patching points.



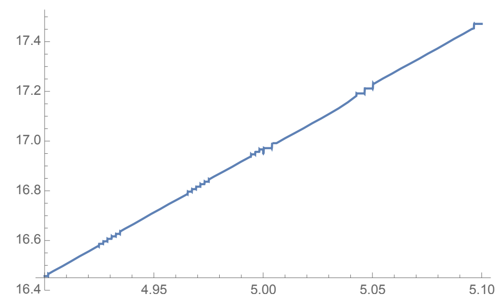
(g) h around $x = 0.462$



(h) h around $x = 1$



(i) h around $x = 2$



(j) h around $x = 5$

4.2 Computation of the likelihood function

We compute the choice probabilities, p_t^l, p_t^l, in closed form using equation (4) from our paper. We then run Monte Carlo simulations to compute the distribution of hitting times F. More precisely, we discretized at the level of decisecond (100 milliseconds). For every combination of parameters we drew 1 million random paths of the discretized Brownian motion (B_t) and for each time compute the first time Z_t exceeds the barrier b^* and the agent stops. From this we obtain a discrete distribution over the times at which the agent stops.

Given p_t^l, p_t^l, F we can compute the likelihood the DDM model assigns to an observation by multiplying the probability of the observed stopping time with the probability of the observed choice conditional on the stopping time.

5 Individual Level Results

5.1 Individual Level Analysis of the Slope of the Boundary

Table 1 lists the parameters (α^*, g^*, h^*) estimated for the unrestricted model $\tilde{b}(t, g, h) = \frac{1}{g+ht}$ and the parameters $(\alpha^\dagger, g^\dagger)$ estimated for the restricted model with $h = 0$. To test the hypothesis $H_0 : (h = 0)$ versus $H_1 : (h > 0)$, we compute the statistic $\eta = -2(O^\dagger - O^*)$, where $O^\dagger$ and O^* are the objective empirical log-likelihoods of the restricted and unrestricted estimation. According to Wilk's Theorem, $\eta \sim \chi^2(1)$ is a chi-squared distributed variable with one degree of freedom. We omit p-values lower than 0.0002.

Table 2 does the same for the model $\check{b}(t, g, h) = g \exp(-ht)$. In both cases, we used the gradient ascent method. There are some entries in those tables in which the difference between the objective values of the restricted and unrestricted models is reported as "-0.0". These correspond to cases where the gradient ascent for the restricted model found a point with larger log-likelihood than the unrestricted model.

Table 3 lists the log likelihoods of both models. Both were fit using the gradient ascent method.

5.2 Individual Level Analysis using the Exact Boundary

Table 4 lists the individual level estimates using a numerically computed b^* function. The 95% confidence intervals are bootstrapped. Table 3 lists the log likelihoods for the optimal boundary. We used the grid search method to estimate this model and for this reason they are not directly comparable to the likelihoods of the approximate and exponential boundaries, which were computed using gradient ascent.

Subject ID	α^*	g^*	h^*	**Wilk's Stat**	p**-value**
	$\alpha^\dagger$	$g^\dagger$		η	$\Pr(\chi^2(1) \geq \eta)$
10	1.7	0.046	0.01297	32.0	
num. obs: 91	3.2	0.148			
11	13.6	0.011	0.00216	54.9	
num. obs: 99	8.7	0.059			
13	4.7	0.053	0.00043	5.2	0.0224
num. obs: 100	3.8	0.080			
14	12.7	0.001	0.00253	101.3	
num. obs: 88	8.1	0.064			
16	12.4	0.018	0.00174	38.0	
num. obs: 100	8.0	0.067			
17	4.5	0.002	0.01227	116.8	
num. obs: 100	3.5	0.165			
18	4.8	0.060	0.00058	10.6	0.0011
num. obs: 100	3.9	0.089			
19	12.6	0.001	0.00267	115.5	
num. obs: 100	9.5	0.056			
20	4.6	0.000	0.01021	98.5	
num. obs: 100	3.2	0.172			
22	13.3	0.003	0.00222	57.1	
num. obs: 100	7.3	0.066			
23	1.8	0.175	0.00014	1.6	0.2105
num. obs: 99	1.8	0.175			
25	4.7	0.070	0.00037	124.4	
num. obs: 99	3.6	0.060			
26	1.6	0.196	0.00048	14.4	
num. obs: 73	3.2	0.143			
27	35.0	0.009	0.00012	10.8	0.0010
num. obs: 100	6.2	0.064			
28	98.2	0.000	0.00042	79.5	
num. obs: 97	8.0	0.070			
29	4.8	0.059	0.00018	5.8	0.0161
num. obs: 100	9.0	0.037			
30	4.7	0.002	0.00219	65.2	
num. obs: 100	3.7	0.071			
31	3.8	0.094	0.00010	−0.0	1.0000
num. obs: 99	3.8	0.094			
32	36.2	0.007	0.00027	27.8	
num. obs: 99	6.1	0.063			

Table 1: Estimated values for the asymptotic boundary - part 1

Subject ID	α^*	g^*	h^*	**Wilk's Stat**	p**-value**
	$\alpha^\dagger$	$g^\dagger$		η	$\Pr(\chi^2(1) \geq \eta)$
33	36.1	0.010	0.00045	12.3	0.0005
num. obs: 49	10.7	0.054			
34	2.0	0.000	0.00743	90.6	
num. obs: 100	1.6	0.143			
35	4.6	0.000	0.00935	95.7	
num. obs: 100	3.2	0.160			
38	4.6	0.000	0.01265	118.1	
num. obs: 100	3.3	0.180			
39	13.4	0.000	0.00232	116.0	
num. obs: 100	8.8	0.057			
40	79.8	0.000	0.00053	113.0	
num. obs: 100	10.6	0.052			
41	4.7	0.060	0.00228	21.5	
num. obs: 100	7.3	0.066			
42	1.6	0.000	0.03928	103.1	
num. obs: 100	1.3	0.347			
44	36.8	0.002	0.00050	60.7	
num. obs: 100	8.0	0.058			
45	2.0	0.137	0.00006	−0.0	1.0000
num. obs: 100	2.0	0.137			
46	4.5	0.047	0.00256	40.7	
num. obs: 98	8.4	0.054			
47	1.9	0.148	0.00511	47.5	
num. obs: 100	3.2	0.148			
48	4.7	0.045	0.00232	38.0	
num. obs: 100	7.2	0.061			
49	4.6	0.000	0.01113	115.9	
num. obs: 100	3.5	0.160			
51	12.5	0.002	0.00235	92.7	
num. obs: 100	9.2	0.054			
52	35.9	0.000	0.00197	97.5	
num. obs: 100	3.4	0.195			
53	14.6	0.010	0.00266	61.1	
num. obs: 100	3.5	0.147			
54	14.6	0.002	0.00259	82.8	
num. obs: 100	7.7	0.069			
55	4.4	0.051	0.00159	29.2	
num. obs: 100	7.3	0.055			
56	4.7	0.047	0.00289	44.4	
num. obs: 100	7.6	0.064			

Table 1: Estimated values for the asymptotic boundary - part 2

Subject ID	α^* $\alpha^\dagger$	g^* $g^\dagger$	h^*	**Wilk's Stat** η	p**-value** $\Pr(\chi^2(1) \geq \eta)$
10	1.7	4.842	0.00223	11.7	0.0006
num. obs: 91	3.2	6.755			
11	4.8	16.370	0.33950	33.2	
num. obs: 99	8.7	16.833			
13	4.6	17.481	0.05060	5.3	0.0212
num. obs: 100	3.8	12.572			
14	4.6	27.651	0.62590	101.4	
num. obs: 88	8.1	15.649			
16	4.7	13.643	0.26032	30.8	
num. obs: 100	8.0	14.981			
17	1.8	6.605	0.60431	82.9	
num. obs: 100	3.5	6.048			
18	4.5	14.667	0.05160	8.5	0.0035
num. obs: 100	3.9	11.226			
19	4.5	29.620	0.72042	124.7	
num. obs: 100	9.5	17.819			
20	1.5	7.133	0.49684	83.1	
num. obs: 100	3.2	5.803			
22	4.5	18.921	0.38403	64.1	
num. obs: 100	7.3	15.252			
23	1.7	5.577	0.01864	3.3	0.0692
num. obs: 99	1.8	5.700			
25	4.6	13.853	0.04496	123.6	
num. obs: 99	3.6	16.681			
26	1.5	4.655	0.00046	11.2	0.0008
num. obs: 73	3.2	7.004			
27	5.0	15.008	0.05665	10.6	0.0011
num. obs: 100	6.2	15.512			
28	4.6	17.427	0.47503	62.7	
num. obs: 97	8.0	14.380			
29	9.0	0.037	0.04543	−0.0	1.0000
num. obs: 100	9.0	26.785			
30	22.3	132.591	0.23887	48.5	
num. obs: 100	3.7	14.016			
31	3.8	0.094	0.04666	−0.0	1.0000
num. obs: 99	3.8	10.629			
32	4.6	14.060	0.05921	20.3	
num. obs: 99	6.1	15.908			

Table 2: Estimated values for the exponential boundary - part 1

Subject ID	α^* $\alpha^\dagger$	g^* $g^\dagger$	h^*	**Wilk's Stat** η	p**-value** $\Pr(\chi^2(1) \geq \eta)$
33	4.6	11.685	0.21512	11.2	0.0008
num. obs: 49	10.7	18.356			
34	4.3	30.101	0.48226	69.4	
num. obs: 100	1.6	7.006			
35	1.9	5.033	0.05710	18.6	
num. obs: 100	3.2	6.243			
38	4.5	39.707	1.55251	149.9	
num. obs: 100	3.3	5.546			
39	4.6	31.002	0.65071	117.2	
num. obs: 100	8.8	17.528			
40	1.6	5.006	0.00046	21.4	
num. obs: 100	10.6	19.402			
41	4.7	14.995	0.23324	2.9	0.0895
num. obs: 100	7.3	15.101			
42	4.6	24.848	1.51480	93.7	
num. obs: 100	1.3	2.882			
44	4.5	18.224	0.23016	45.0	
num. obs: 100	8.0	17.353			
45	2.0	0.137	0.05375	-0.0	1.0000
num. obs: 100	2.0	7.291			
46	4.5	15.503	0.19851	35.2	
num. obs: 98	8.4	18.422			
47	4.6	14.300	0.33069	35.0	
num. obs: 100	3.2	6.757			
48	4.6	15.625	0.17315	31.4	
num. obs: 100	7.2	16.324			
49	4.5	34.380	1.25724	129.8	
num. obs: 100	3.5	6.238			
51	4.6	23.833	0.51900	87.6	
num. obs: 100	9.2	18.577			
52	21.4	96.043	1.05631	84.9	
num. obs: 100	3.4	5.136			
53	4.6	12.740	0.30209	37.2	
num. obs: 100	3.5	6.818			
54	4.7	16.586	0.34248	53.9	
num. obs: 100	7.7	14.568			
55	4.7	13.572	0.05377	18.9	
num. obs: 100	7.3	18.178			
56	4.6	15.662	0.23117	38.0	
num. obs: 100	7.6	15.704			

Table 2: Estimated values for the exponential boundary - part 2

Subject ID	O^*_{asym}	O^*_{exp}	O^*_{opt}
10	−309.8423	−319.9846	−310.969
11	−331.1384	−342.0040	−254.803
13	−454.3407	−454.2937	−451.762
14	−276.5247	−276.4654	−307.022
16	−345.0759	−348.6462	−347.601
17	−276.0078	−292.9507	−309.468
18	−416.0104	−417.0699	−414.433
19	−310.0567	−305.4688	−344.884
20	−291.5485	−299.2674	−318.67
22	−351.0997	−347.6186	−360.005
23	−382.9425	−382.0749	−377.709
25	−423.7458	−424.1767	−415.441
26	−253.3470	−254.9518	−247.451
27	−420.2934	−420.3970	−417.718
28	−306.9961	−315.3583	−326.716
29	−450.1698	−453.0670	−447.866
30	−399.7588	−408.0832	−349.164
31	−411.3307	−411.3307	−403.622
32	−381.6903	−385.4122	−307.389
33	−165.9013	−166.4507	−149.061
34	−330.1436	−340.7326	−365.663
35	−292.3161	−330.8427	−320.064
38	−267.8982	−252.0136	−303.2
39	−318.1265	−317.4799	−354.198
40	−298.4855	−344.2907	−329.924
41	−371.1066	−380.4350	−369.735
42	−247.0126	−251.7067	−278.896
44	−367.3603	−375.1763	−381.01
45	−414.8559	−414.8559	−408.797
46	−367.4805	−370.2244	−370.454
47	−335.9341	−342.1610	−342.249
48	−383.0223	−386.3268	−385.92
49	−277.0366	−270.0774	−312.498
51	−333.0213	−335.5746	−358.194
52	−274.2847	−280.5827	−299.424
53	−317.7696	−329.7346	−331.691
54	−314.3696	−328.8399	−222.119
55	−397.8434	−402.9856	−395.557
56	−369.5025	−372.6943	−376.387

Table 3: Log likelihoods for the asymptotic and exponential models (using gradient ascent method, columns 1 and 2) and for the optimal estimated boundary (using grid search method, column 3).

Subject ID	α	σ	c
10	1.8	1.3	0.05
num. obs: 91	(1.6, 2.4)	(0.5, 1.8)	(0.02, 0.08)
11	5.2	4.8	0.37
num. obs: 76	(4, 6)	(3.1, 8)	(0.22, 0.6)
13	2.8	1.1	0.02
num. obs: 100	(2.4, 6)	(0.6, 2.5)	(0.01, 0.06)
14	5	4.7	0.3
num. obs: 88	(2.6, 6)	(2.3, 6.8)	(0.12, 0.39)
16	2.8	1.4	0.08
num. obs: 100	(2, 4.8)	(1.1, 4)	(0.05, 0.27)
17	5.6	8.1	0.77
num. obs: 100	(4.8, 6)	(4.2, 9.9)	(0.43, 0.89)
18	3.8	1.6	0.05
num. obs: 100	(2.4, 5.4)	(0.7, 2.7)	(0.02, 0.09)
19	4.4	4.3	0.27
num. obs: 100	(2.8, 6)	(1.9, 7.1)	(0.1, 0.43)
20	2	1.8	0.1
num. obs: 100	(1.4, 5.4)	(0.7, 8.3)	(0.03, 0.59)
22	3	2.3	0.1
num. obs: 100	(2.2, 5.4)	(1.3, 4.9)	(0.05, 0.27)
23	2	0.6	0.02
num. obs: 99	(1.4, 2.6)	(0.4, 2.8)	(0.01, 0.05)
25	4	0.5	0.01
num. obs: 99	(1.8, 6)	(0.4, 1.1)	(0.01, 0.03)
26	1.8	1.1	0.04
num. obs: 73	(1.4, 2.2)	(0.7, 1.1)	(0.02, 0.05)
27	5.6	2.5	0.1
num. obs: 100	(3, 6)	(1.6, 6)	(0.05, 0.15)
28	5.4	6.6	0.43
num. obs: 97	(3, 6)	(2.4, 7.5)	(0.15, 0.5)
29	3.8	0.9	0.02
num. obs: 100	(3.6, 6)	(0.6, 2.7)	(0.01, 0.05)
30	3.8	1.3	0.03
num. obs: 82	(2.6, 6)	(1.2, 3.2)	(0.02, 0.09)
31	2	1.4	0.02
num. obs: 99	(1.2, 2.2)	(0.4, 2.3)	(0.01, 0.02)
32	5	3.1	0.11
num. obs: 78	(2, 6)	(1, 6.2)	(0.02, 0.25)

Table 4: Estimated values for the exact boundary - part 1. The numbers in intervals are bootstrapped 95% confidence intervals.

Subject ID	α	σ	c
33	3.4	2.6	0.14
num. obs: 44	(1.6, 5.8)	(1, 6.7)	(0.03, 0.52)
34	2.2	1	0.03
num. obs: 100	(1.6, 2.2)	(0.5, 1)	(0.01, 0.03)
35	2.2	1.7	0.09
num. obs: 100	(1.8, 4.4)	(1.4, 4.5)	(0.06, 0.41)
38	3.2	4.3	0.39
num. obs: 100	(1.6, 6)	(1.1, 9.6)	(0.07, 0.94)
39	5.8	5.8	0.32
num. obs: 100	(4.2, 6)	(3.8, 8.1)	(0.21, 0.41)
40	2.2	1.7	0.08
num. obs: 100	(1.8, 6)	(0.9, 6.6)	(0.03, 0.52)
41	3.4	1.5	0.07
num. obs: 100	(2.2, 5.4)	(0.6, 3.6)	(0.02, 0.21)
42	1.2	1.1	0.06
num. obs: 100	(1.2, 4.6)	(0.9, 6.6)	(0.06, 0.89)
44	5.4	3.8	0.19
num. obs: 100	(2, 6)	(0.6, 4.5)	(0.02, 0.24)
45	2	1.8	0.02
num. obs: 100	(1.6, 3)	(0.5, 1.9)	(0.01, 0.04)
46	3	2.1	0.08
num. obs: 98	(1.8, 5)	(1.2, 2.7)	(0.03, 0.13)
47	2	1	0.06
num. obs: 100	(1.4, 3)	(0.8, 2.7)	(0.04, 0.16)
48	3	2.1	0.08
num. obs: 100	(1.4, 3.8)	(0.6, 2.5)	(0.02, 0.12)
49	4.4	4.5	0.41
num. obs: 100	(1.4, 6)	(0.7, 9.5)	(0.03, 0.86)
51	6	5.5	0.32
num. obs: 100	(1.6, 6)	(0.7, 5.5)	(0.03, 0.32)
52	5.6	8.5	0.88
num. obs: 100	(1.6, 6)	(1.7, 8.5)	(0.11, 0.89)
53	5.8	5.1	0.44
num. obs: 100	(3, 6)	(2.6, 8.4)	(0.2, 0.68)
54	5.8	5.7	0.47
num. obs: 66	(3.2, 5.8)	(1.5, 7)	(0.09, 0.53)
55	2.4	1.1	0.03
num. obs: 100	(1.8, 3.4)	(0.5, 2.3)	(0.01, 0.06)
56	3.6	2.4	0.13
num. obs: 100	(1.6, 6)	(0.04, 0.27)	(0.8, 4.7)

Table 4: Estimated values for the exact boundary - part 2. The numbers in intervals are bootstrapped 95% confidence intervals.

Subject ID	Correct	Incorrect	t-statistic	p-value
10	17.79	16.67	0.46	
11	16.38	17.69	−0.58	
13	45.30	41.81	0.44	
14	18.77	18.96	−0.17	
16	17.06	18.41	−0.64	
17	12.46	13.51	−1.38	
18	30.37	41.37	−2.37	0.0237
19	17.86	18.77	−0.75	
20	14.15	13.79	0.22	
22	18.21	30.40	−4.75	0.0001
23	22.45	39.02	−1.84	0.0871
25	30.37	42.18	−1.28	
26	14.69	29.86	−2.84	0.0132
27	32.10	31.40	0.17	
28	16.24	15.08	0.75	
29	39.76	60.30	−1.77	0.0990
30	36.79	43.22	−1.41	
31	30.50	47.38	−2.02	0.0590
32	24.01	36.00	−2.73	0.0147
33	15.49	13.39	0.86	
34	22.92	30.27	−2.78	0.0191
35	14.22	15.08	−0.61	
38	11.70	14.69	−3.11	0.0058
39	18.69	22.10	−2.25	0.0312
40	15.15	18.19	−1.99	0.0573
41	18.71	28.72	−2.07	0.0522
42	9.57	9.58	−0.02	
44	25.13	23.18	0.84	
45	29.40	46.33	−2.10	0.0481
46	23.56	24.88	−0.24	
47	16.56	20.58	−1.74	0.0915
48	24.84	30.00	−1.30	
49	12.81	14.51	−1.44	
51	19.00	25.81	−2.92	0.0084
52	10.44	15.10	−3.09	0.0058
53	14.10	18.83	−1.94	0.0665
54	14.58	20.53	−3.08	0.0043
55	27.02	36.81	−1.60	
56	21.47	27.44	−2.05	0.0488

Table 5: Mean response times (0.1s) for correct and incorrect decisions. The third and fourth columns test the null that the mean decision time for correct and incorrect are equal, using the two-sample Welch t-test. p-values that exceeds 10% are left out.

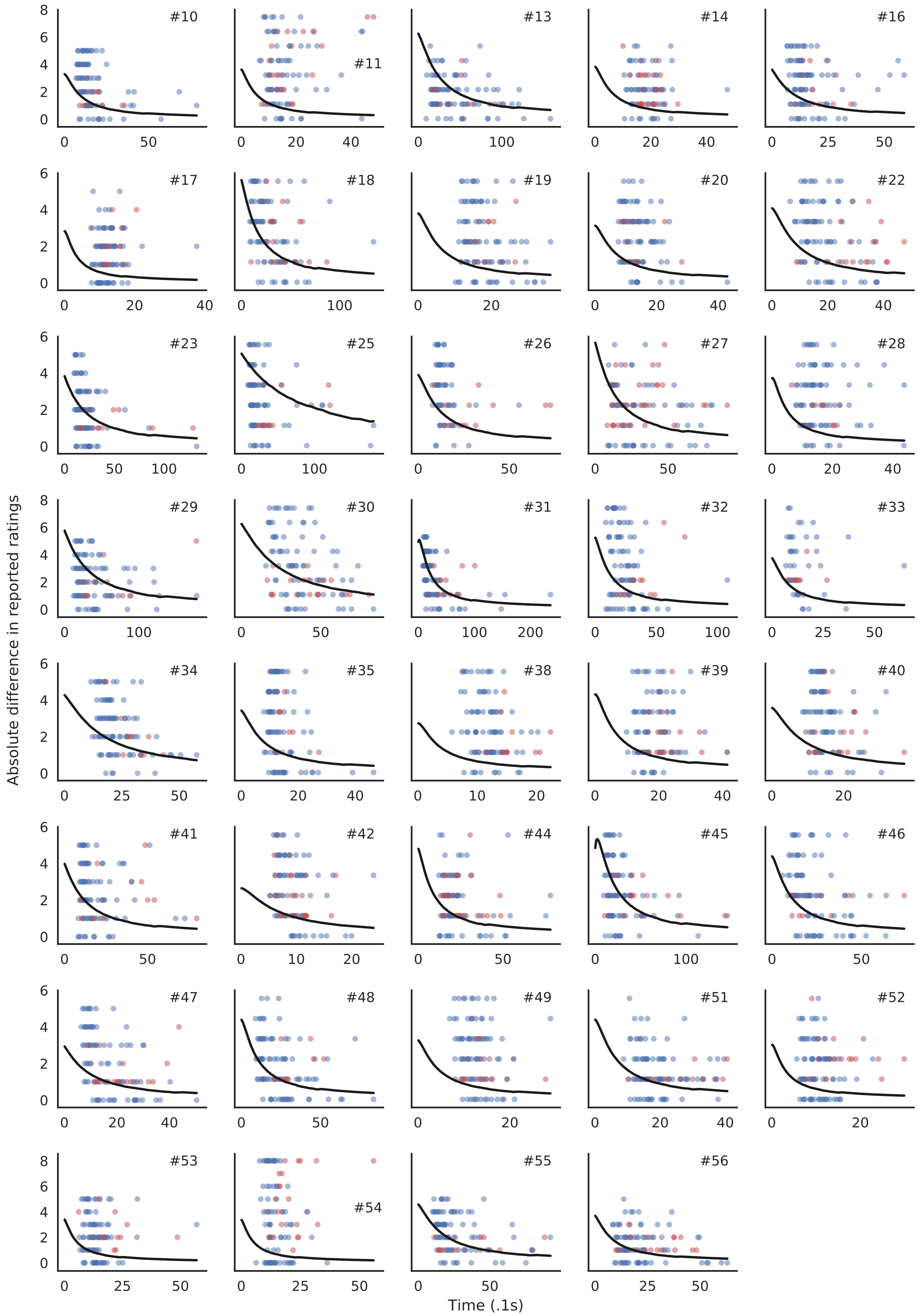


Figure 3: Estimated optimal boundaries for different subjects. Blue dots are correct decisions, red dots are incorrect ones.

References

Bather, John A. 1962. "Bayes procedures for deciding the sign of a normal mean." Vol. 58, 599–620, Cambridge Univ Press.

Dudley, Richard M. 2002. *Real analysis and probability.* Vol. 74, Cambridge University Press.

Kruse, Thomas, and Philipp Strack. 2015. "Optimal stopping with private information." *forthcoming in Journal of Economic Theory.*

Revuz, Daniel, and Marc Yor. 1999. *Continuous martingales and Brownian motion.* Vol. 293, Springer.