

Online Appendix to Stochastic Choice and Optimal Sequential Sampling

Drew Fudenberg* Philipp Strack[†] Tomasz Strzalecki[‡]

March 16, 2017

1 Proofs omitted in the main text

1.1 Proof of Lemma 4

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

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

*Department of Economics, MIT

[†]Department of Economics, University of California, Berkeley

[‡]Department of Economics, Harvard University

Proof of (14)

By Lemma 3, $V(0, x^l, x^r, c\lambda, \sigma_0, \alpha)$ equals

$$\begin{aligned}
& \lambda^{-1} \sup_{\tau'} \mathbb{E} \left[\max \left\{ \lambda x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dM_s^l, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dM_s^r \right\} - c\tau' \right] \\
& \lambda^{-1} \lambda^2 \sup_{\tau'} \mathbb{E} \left[\lambda^{-2} \max \left\{ \lambda x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dM_s^l, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dM_s^r \right\} - c\tau' \lambda^{-2} \right] \\
& = \lambda \sup_{\tau'} \mathbb{E} \left[\max \left\{ \lambda^{-1} x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\lambda^2 \sigma_0^{-2} + s\alpha^{-2}} dM_s^l, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\lambda^2 \sigma_0^{-2} + s\alpha^{-2}} dM_s^r \right\} - c\tau' \lambda^{-2} \right] \\
& = \lambda V(0, x_1 \lambda^{-1}, x_2 \lambda^{-1}, c\lambda^{-2}, \sigma_0/\lambda, \alpha).
\end{aligned}$$

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

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

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

$$\begin{aligned}
& = \sup_{\tau \geq t} \mathbb{E} \left[\max \left\{ x_1 + \int_t^\tau \frac{\lambda^{-1}\alpha^{-1}}{\sigma_0^{-2} + \lambda^{-2}\alpha^{-2}s} dW_s^1, x_2 + \int_t^\tau \frac{\lambda^{-1}\alpha^{-1}}{\sigma_0^{-2} + \lambda^{-2}\alpha^{-2}s} dW_s^2 \right\} - c(\tau - t) \right] \\
& = \lambda \sup_{\tau \geq t} \mathbb{E} \left[\max \left\{ \lambda^{-1}x_1 + \int_t^\tau \frac{\alpha^{-1}}{\lambda^2\sigma_0^{-2} + \alpha^{-2}s} dW_s^1, \lambda^{-1}x_2 + \int_t^\tau \frac{\alpha^{-1}}{\lambda^2\sigma_0^{-2} + \alpha^{-2}s} dW_s^2 \right\} - (c\lambda^{-1})(\tau - t) \right] \\
& = \lambda V(t, \lambda^{-1}x_1, \lambda^{-1}x_2, \lambda^{-1}c, \lambda^{-1}\sigma_0, \alpha).
\end{aligned}$$

Thus,

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

Proof of (16)

By Lemma 3, $V(0, x^l, x^r, c\lambda, \sigma_0, \alpha)$ equals

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

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

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

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

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

1.2 Proof of Proposition 4

We rely on [John A 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 $\hat{X}_t^i = t^{-1} Z_t^i$ and $\hat{\sigma}_t^{-2} = t\alpha^{-2}$. In Section 6, p. 619 [Bather \(1962\)](#) shows that

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

which implies that

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

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,

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

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

$$\left|k(t, c, \infty, \alpha) - \frac{1}{2c\alpha^{-2}t^2}\right| \leq \frac{\beta}{t^{7/2}}.$$

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 $k(t, c, \infty, \alpha) = k(s, c, \sigma_0, \alpha)$, so

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

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

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

1.3 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 $t = \alpha^2 \tilde{\sigma}_0^{-2} (c^{2/3} - 1)$, we have $\tilde{\sigma}_t = \tilde{\sigma}_0 c^{-1/3}$ and using equation (13) yields

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

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

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

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

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 $\beta, \underline{c} > 0$ independent of c such that for all $c \geq \underline{c}$

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

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

1.4 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

$$\begin{aligned} \sup_{\tau} \mathbb{E} [\max\{X_{\tau}^l, X_{\tau}^r\} - d(\tau)] &= \sup_{\tau} \mathbb{E} \left[\frac{1}{2} \max\{X_{\tau}^l - X_{\tau}^r, X_{\tau}^r - X_{\tau}^l\} + \frac{1}{2} (X_{\tau}^l + X_{\tau}^r) - d(\tau) \right] \\ &= \sup_{\tau} \mathbb{E} \left[\frac{1}{2} |X_{\tau}^l - X_{\tau}^r| - d(\tau) \right] + \frac{1}{2} (X_0^l + X_0^r), \end{aligned}$$

so any optimal stopping time also solves $\sup_{\tau} \mathbb{E} [|X_{\tau}^l - X_{\tau}^r| - 2d(\tau)]$. Define $\Delta_n = |X_{t_n}^l - X_{t_n}^r|$ for all $n = 1, \dots, N$. Observe that $(\Delta_n)_{n=1, \dots, N}$ 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 [Thomas Kruse and Philipp 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 .

1.4.1 Certain-Difference DDM

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

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

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

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

Let $\psi_n := [Z_{t_n}^l - Z_{t_n}^r] - [Z_{t_{n-1}}^l - Z_{t_{n-1}}^r]$ 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

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

(1): It is easily seen that $\mathbb{E}[\Delta_{n+1} \mid \mathcal{F}_{t_n}] \leq (\theta'' - \theta')$, so for C big enough, $\mathbb{E}[\Delta_{n+1} \mid \mathcal{F}_{t_n}] \leq C(1 + \Delta_n)$.

(2): To simplify notation we introduce $m_n = |l_n|$. The process $(m_n)_{n=1, \dots, N}$ 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

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

and variance

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

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 (2) increases in m_n it follows that m_{n+1} increases in m_n in the sense of first order stochastic dominance. By (1) $m_n = |l_n|$ 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 $(p_n)_{n=1, \dots, M}$ is a martingale, and moreover conditioning on p is equivalent to conditioning on $1 - p$, we have that

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

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

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

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 .

1.4.2 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

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

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

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

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

(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

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

As $\Delta_i = |X_{t_i}^l - X_{t_i}^r|$ 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

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

Hence, z is strictly decreasing in y . □

2 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).

2.1 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; all the proofs are contained in Section ??.

Lemma 1.

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

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. $\square$

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 2.

$$V(t, x, 0, 1, \sigma_0, 1) = \sup_{\tau \geq \psi(t)} \mathbb{E} [\max\{W_\tau, 0\} - (q(\tau) - q(\psi(t))) \mid W_{\psi(t)} = x], \quad (3)$$

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 x_t of X_t . That is, $x_{t+\Delta_t}$ takes values $x_t - \Delta_x$ and $x_t + \Delta_x$ with equal probabilities. The discrete time version of the optimization problem is:

$$v(t, x, 0, 1, \sigma_0, 1) = \max \left\{ x, \frac{1}{2}v(t + \Delta_t, x + \Delta_x) + \frac{1}{2}v(t + \Delta_t, x - \Delta_x) - (q(t + 1) - q(t)) \right\} \quad (4)$$

The solution to (4) approaches the solution to (3) when $\Delta_t, \Delta_x \rightarrow 0$, as long as $\Delta_t < \Delta_x^2$.

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.

2.1.1 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.

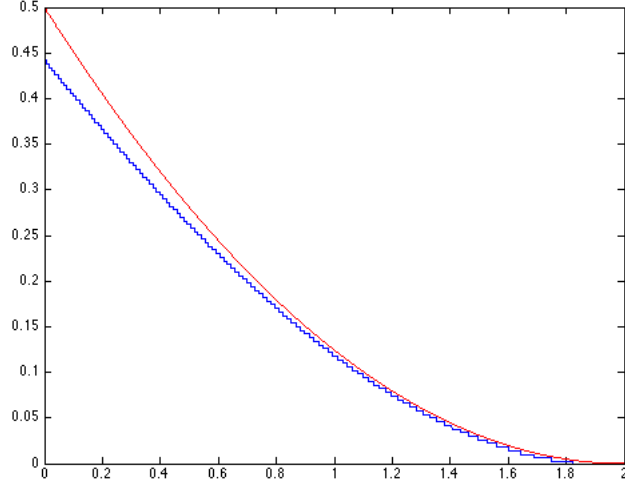


Figure 1: The function $\check{k}$ computed for $\sigma_0 = 1$ (blue) and the function $\bar{k}$ (red).

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 $\psi(s) = se^{-\frac{1}{2}s^2} \int_0^s e^{\frac{1}{2}y^2} dy = \sqrt{\frac{\pi}{2}}se^{-\frac{1}{2}s^2}\text{erfi}\left(\frac{s}{\sqrt{2}}\right)$, 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

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

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:

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

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.

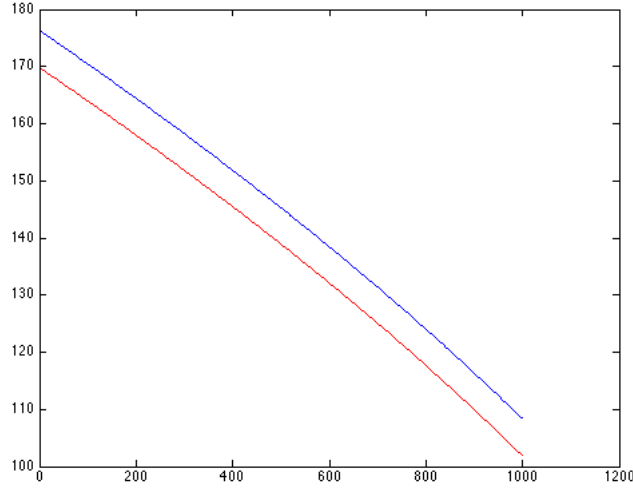


Figure 2: The function $\tilde{k}$ computed for $\sigma_0 = 30$ (blue) and the implicit function $s \mapsto x$ (red).

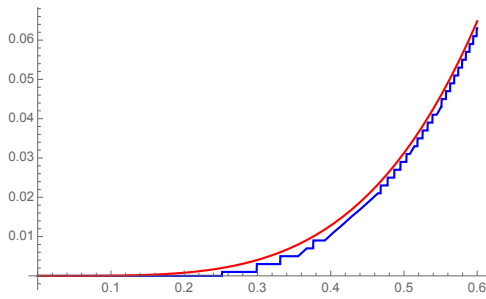
2.1.2 The function h used to compute the likelihood function

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 $\bar{h}(x) := \bar{k}(0, 1, x, 1) = 0.5x^4$.

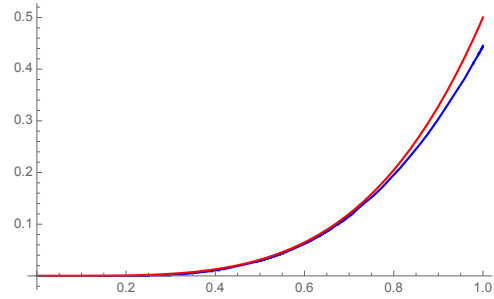
We compute the function h as follows:

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

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



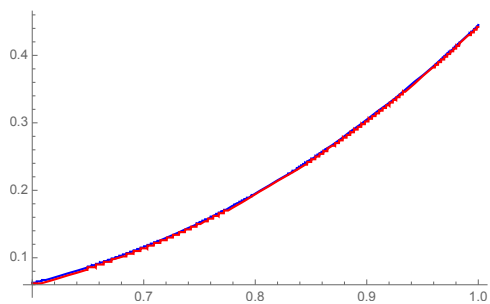
(a) $\bar{h}$ (red) and h_1 (blue) for $x \in [0, .6)$



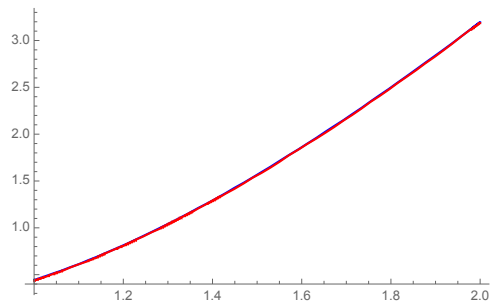
(b) $\bar{h}$ (red) and h_1 (blue) for $x \in [0, 1)$

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 .

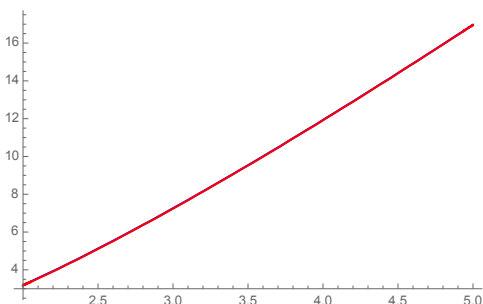
The following graph shows that the consecutive h_{σ_0} functions lie pretty much on top of each other.



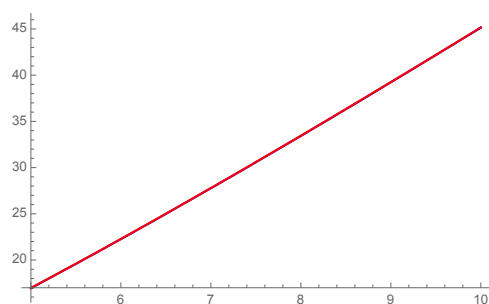
(c) h_1 (red) and h_2 (blue) for $x \in [0.6, 1]$



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

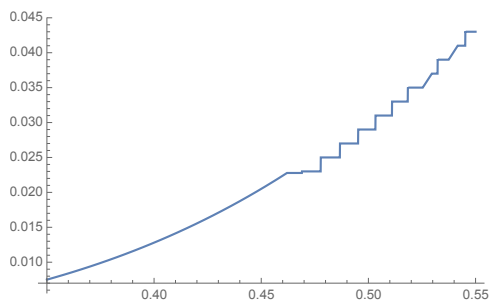


(e) h_5 (red) and h_{10} (blue) for $x \in [2, 5]$

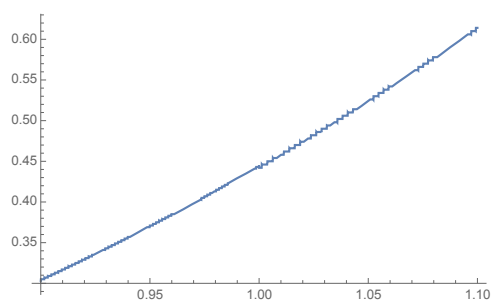


(f) h_{10} (red) and h_{30} (blue) for $x \in [5, 10]$

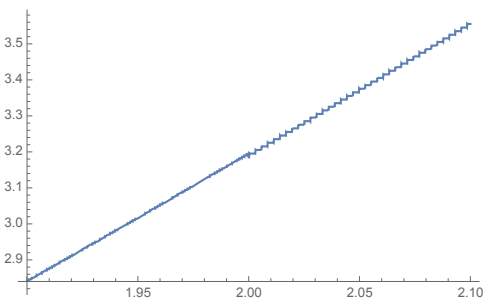
The following graph displays the h function around the patching points.



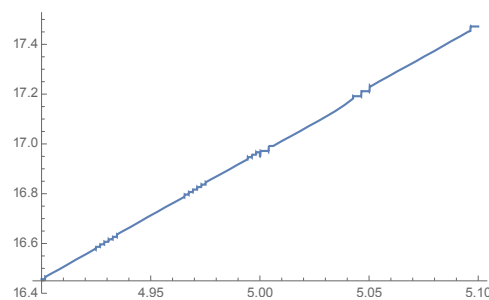
(g) h around $x = 0.462$



(h) h around $x = 1$



(i) h around $x = 2$



(j) h around $x = 5$

2.2 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.

3 Individual Level Results

3.1 Individual Level Analysis of the Slope of the Boundary

Table 1 lists the parameters (α^*, g^*, h^*) estimated for the unrestricted model 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.

3.2 Individual Level Analysis using the Exact Boundary

Table 2 lists the individual level estimates using a numerically computed b function. The standard errors are bootstrapped.

Subject ID	α^* $\alpha^\dagger$	a^* $a^\dagger$	d^*	Wilk's Stat η	p -value $\Pr(\chi^2(1) \geq \eta)$
10	2.2	0.12	0.065	21.7	
<i>num. obs:</i> 91	2.8	0.21			
11	5.0	0.01	0.085	51.8	
<i>num. obs:</i> 76	4.2	0.14			
13	3.0	0.09	0.0033	6.0	0.0145
<i>num. obs:</i> 100	4.4	0.08			
14	3.2	0.01	0.095	87.4	
<i>num. obs:</i> 88	2.6	0.19			
16	3.4	0.06	0.075	33.9	
<i>num. obs:</i> 100	3.4	0.16			
17	4.0	0.01	0.145	106.8	
<i>num. obs:</i> 100	2.2	0.28			
18	3.6	0.08	0.01	13.1	0.0003
<i>num. obs:</i> 100	3.2	0.11			
19	3.4	0.01	0.095	103.36	
<i>num. obs:</i> 100	2.6	0.19			
20	2.6	0.01	0.205	88.6	
<i>num. obs:</i> 100	2.0	0.24			
22	3.0	0.04	0.07	49.1	
<i>num. obs:</i> 100	2.6	0.18			
23	2.6	0.12	0.03	7.4	0.0064
<i>num. obs:</i> 99	2.6	0.18			
25	3.8	0.12	0.0	0.0	1.0
<i>num. obs:</i> 99	3.8	0.12			
26	2.6	0.08	0.08	22.7	
<i>num. obs:</i> 73	2.4	0.2			
27	5.0	0.07	0.0049	9.2	0.0024
<i>num. obs:</i> 100	5.0	0.08			
28	4.6	0.01	0.095	68.2	
<i>num. obs:</i> 97	3.2	0.17			
29	4.4	0.07	0.0027	6.6	0.0103
<i>num. obs:</i> 100	4.6	0.08			
30	4.6	0.01	0.02	48.2	
<i>num. obs:</i> 82	3.2	0.09			
31	2.8	0.16	0.0002	0.6	0.4318
<i>num. obs:</i> 99	2.8	0.16			
32	5.0	0.04	0.02	24.4	
<i>num. obs:</i> 78	4.6	0.1			

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

Subject ID	α^* $\alpha^\dagger$	a^* $a^\dagger$	d^*	Wilk's Stat η	p -value $\Pr(\chi^2(1) \geq \eta)$
33	3.2	0.04	0.12	13.4	0.0002
<i>num. obs:</i> 44	4.2	0.14			
34	2.6	0.01	0.075	87.0	
<i>num. obs:</i> 100	2.2	0.15			
35	2.8	0.01	0.18	92.6	
<i>num. obs:</i> 100	2.2	0.22			
38	2.8	0.01	0.255	108.2	
<i>num. obs:</i> 100	2.0	0.26			
39	4.2	0.01	0.07	100.2	
<i>num. obs:</i> 100	3.0	0.14			
40	2.6	0.01	0.145	94.9	
<i>num. obs:</i> 100	2.4	0.2			
41	4.2	0.08	0.02	14.3	0.0002
<i>num. obs:</i> 100	3.6	0.15			
42	2.6	0.01	0.415	96.2	
<i>num. obs:</i> 100	1.2	0.43			
44	4.6	0.02	0.04	58.4	
<i>num. obs:</i> 100	3.4	0.14			
45	2.2	0.14	0.0044	3.4	0.0643
<i>num. obs:</i> 100	2.2	0.15			
46	3.4	0.07	0.035	30.5	
<i>num. obs:</i> 98	3.4	0.14			
47	2.4	0.09	0.115	44.9	
<i>num. obs:</i> 100	2.6	0.23			
48	2.6	0.1	0.045	32.0	
<i>num. obs:</i> 100	2.6	0.18			
49	2.4	0.01	0.22	104.1	
<i>num. obs:</i> 100	2.6	0.25			
51	2.8	0.01	0.135	83.3	
<i>num. obs:</i> 100	2.2	0.22			
52	3.2	0.01	0.24	85.0	
<i>num. obs:</i> 100	2.8	0.31			
53	4.6	0.02	0.105	55.2	
<i>num. obs:</i> 100	3.4	0.18			
54	4.2	0.01	0.1	53.3	
<i>num. obs:</i> 66	3.2	0.17			
55	2.8	0.1	0.025	18.9	
<i>num. obs:</i> 100	2.2	0.15			
56	3.6	0.07	0.045	34.9	
<i>num. obs:</i> 100	3.2	0.15			

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

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

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

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

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

References

- Bather, John A.** 1962. “Bayes procedures for deciding the sign of a normal mean.” Vol. 58, 599–620, Cambridge Univ Press.
- Kruse, Thomas, and Philipp Strack.** 2015. “Optimal stopping with private information.” *forthcoming in Journal of Economic Theory*.