

# Switching Cost Models as Hypothesis Tests\*

Samuel N. Cohen  
University of Oxford<sup>†</sup>

Timo Henckel  
Australian National University & CAMA

Gordon D. Menzies  
University of Technology Sydney & CAMA<sup>‡</sup>

Johannes Muhle-Karbe  
Carnegie Mellon University

Daniel J. Zizzo  
Newcastle University & CAMA

August 27, 2018

## Abstract

We relate models based on costs of switching beliefs (e.g. due to inattention) to hypothesis tests. Specifically, for an inference problem with a penalty for mistakes and for switching the inferred value, a band of inaction is optimal. We show this band is equivalent to a confidence interval, and therefore to a two-sided hypothesis test.

*Keywords:* inference; switching cost; inferential expectations, hypothesis test.

*JEL classification codes:* D01, D81, D84

## 1 Introduction

This paper provides a new micro-foundation for two-sided hypothesis tests. Agents receive sequential information and conduct inference which penalizes adjustments to the estimator and deviations from the classical Bayesian estimate. We show that, to a first order approximation for small adjustment costs, the resulting estimator has a band of inaction with width proportional to the Bayesian estimator's standard deviation. This makes it equivalent to a confidence interval and therefore to a two-sided hypothesis test.

---

\*We acknowledge helpful comments of colleagues at the University of Technology Sydney.

<sup>†</sup>Research supported by the Oxford–Man Institute for Quantitative Finance

<sup>‡</sup>Gordon D. Menzies (corresponding author), School of Economics and Finance, University of Technology Sydney, Broadway, Sydney NSW 2007, Australia. E-mail address: gordon.menzies@uts.edu.au. Phone: +61-2-95147728.

Our result locates belief formation models based on hypothesis tests, such as Menzies and Zizzo’s (2009) inferential expectations model, within a wider literature on switching costs due to sticky belief adjustment. Switching costs may arise, say, from ‘menu costs’, transactions in illiquid markets, cognitive effort in attention and observation or the consultation of experts (Caplin and Spulber, 1987; Alvarez et al., 2017; Magnani et al., 2016; Carroll, 2003).

State-dependent belief adjustments describe how new information about the underlying economic state  $X_t$  is incorporated. In our model, agents passively observe until new information exceeds a threshold, depending on the uncertainty of the estimated state, and only then readjust their policy. This infrequent adjustment is similar to models of inattention and portfolio choice (Abel et al., 2013; Huang and Liu, 2007).

A concrete example for our analysis is portfolio choice with partial information. Here,  $X_t$  represents the unknown expected returns which are estimated from time-series data. If  $\hat{X}_t$  denotes the Bayesian estimate of  $X_t$ , the optimal portfolio then is typically of the form  $h(\hat{X}_t)$ . With transaction costs, this ideal portfolio cannot be implemented and instead has to be replaced by an approximation  $h(\Theta_t)$ , where  $\Theta_t$  is an alternative estimate of  $X_t$  that only changes infrequently. The optimal  $\Theta_t$  is in turn identified by our tradeoff between switching costs and inefficiency costs due to deviations from the optimal estimator.

Key to our approach is the use of asymptotic approximation methods, to allow closed-form solutions valid when costs are small. This typically yields an approximate ‘no-action region’, within which agents accept deviations from the no-cost optimum (Korn, 1998; Lo et al., 2004). Our specific contribution is to link switching cost models to hypothesis tests using the results of Altarovici et al. (2015). Their purpose is to describe trade within a financial market; we propose that their asymptotic approximation can also be applied to a wide class of recursive estimation problems.

An appendix contains an illustration of our model in discrete time.

## 2 The model

We base our setting on a Kalman–Bucy filter (Kalman and Bucy, 1961), as this has a wide variety of applications (see Bain and Crisan, 2009).

We write  $X$  for a multivariate (hidden) process, which we seek to estimate using multivariate observations  $Y$ . We suppose  $X$  and  $Y$  satisfy

$$\begin{cases} dX_t = F_t X_t dt + dW_t, & X_0 \sim N(\hat{X}_0, P_0), \\ dY_t = A_t X_t dt + dB_t, & Y_0 = 0, \end{cases}$$

where  $W$  and  $B$  are independent continuous martingales, with quadratic variations

$$d\langle W \rangle_t = Q_t dt, \quad d\langle B \rangle_t = R_t dt.$$

Here  $F, A, Q$  and  $R$  are matrix-valued deterministic functions of appropriate dimensions,  $R$  is invertible and  $A$  is nonzero, and  $(\hat{X}_0, P_0)$  are the mean and

variance of our initial estimate of  $X_0$ . The filtration  $\mathcal{F}_t = \sigma(Y_s; s \leq t)$  represents the information available from observing  $Y$  up to time  $t$ .

For these dynamics, conditional on our observations  $\{Y_s\}_{0 \leq s \leq t}$ , the hidden state  $X_t$  has a multivariate normal distribution:

$$X_t | \mathcal{F}_t \sim N(\hat{X}_t, P_t).$$

The values of  $(\hat{X}_t, P_t)$  have joint dynamics

$$\begin{cases} d\hat{X}_t = F_t \hat{X}_t dt + K_t d\hat{V}_t, \\ dP_t/dt = F_t P_t + P_t F_t^\top + Q_t - K_t R_t K_t^\top, \end{cases} \quad (1)$$

with initial values  $(\hat{X}_0, P_0)$ , where  $K_t = P_t A_t^\top R_t^{-1}$  denotes the *Kalman gain* process, and  $d\hat{V}_t = dY_t - A_t \hat{X}_t dt$  defines the *innovations process*  $\hat{V}$ , which is a martingale under  $\{\mathcal{F}_t\}_{t \geq 0}$ , with quadratic variation  $d\langle \hat{V} \rangle_t = R_t dt$ .

*Example 1.* A simple example is when our processes are all scalar,  $F, Q \equiv 0$  and  $A, R \equiv 1$ . Then  $X \equiv X_0$  is a (unknown) constant and  $K_t = P_t$ , so

$$\begin{aligned} \frac{dP_t}{dt} = -K_t R_t K_t^\top = -P_t^2 &\quad \Rightarrow \quad P_t = \frac{1}{1/P_0 + t}, \\ d\hat{X}_t = K_t d\hat{V}_t = \frac{1}{1/P_0 + t} d\hat{V}_t. \end{aligned}$$

Observe that the posterior variance  $P_t$  collapses like  $1/t$ , as we would expect from a standard observation problem. This is the continuous-time analogue of a Bayesian estimation problem for an unknown mean with normal errors, with prior  $N(\hat{X}_0, P_0)$  leading to posterior  $N(\hat{X}_t, P_t)$ .

We suppose that, over a fixed time period  $[0, T]$ , our agent estimates  $X_t$  with an approximation  $\Theta_t$  of  $\hat{X}_t$ . She has initial wealth  $z$ , from which she continuously pays monetary costs  $\rho(\hat{X}_t - \Theta_t)$  due to tracking error relative to the optimal filter estimate and a cost  $\lambda$  whenever  $\Theta_t$  changes. We assume  $\rho$  is convex, smooth and minimized at  $\rho(0) = 0$ . For a utility function  $U$ , our agent wishes to optimize her utility of expected wealth

$$J(\omega, t, z, \Theta; \lambda) = E\left[U(Z_T)\right] = E\left[U\left(z - \int_t^T \rho(\hat{X}_t - \Theta_t) - \lambda \sum_{t \leq s \leq T} I_{\{\Delta \theta_s \neq 0\}}\right) \middle| \mathcal{F}_t\right]$$

over piecewise constant adapted processes  $\Theta$ . As  $\hat{X}$  is a Markov process, there exists a value function

$$v(t, \hat{X}_t(\omega), z, \Theta_t; \lambda) = \sup_{\Theta': \Theta_t = \Theta_t'} J(\omega, t, z, \Theta'; \lambda).$$

Like in Korn (1998), Lo et al. (2004) and Altarovici et al. (2015), the value function can be expanded<sup>1</sup> in powers of  $\lambda$ . If  $\lambda$  is small, by ignoring higher order terms, we obtain an approximation to  $v$ , and hence to the optimal choice of  $\Theta$ .

<sup>1</sup>Corresponding verification theorems could be derived using stability results for viscosity solutions as in Altarovici et al. (2015) or martingale methods, cf. Feodorina (2016).

### 3 Dynamic programming

With fixed adjustment costs, it will be optimal to leave  $\Theta$  unchanged until  $\hat{X}_t - \Theta_t$  is sufficiently large. Write  $\mathfrak{R}$  for the region where  $\Theta$  remains fixed. A standard dynamic programming argument yields a partial differential equation for the value function  $v(t, \hat{x}, z, \theta; \lambda)$ .

Indeed, the optimal filter  $\hat{X}$  without adjustment costs has the diffusive dynamics (1). By the martingale optimality principle, the value function evaluated along the state variables  $(t, \hat{X}_t, Z_t, \Theta_t)$  is a martingale for the optimal  $\Theta$ , and a supermartingale otherwise. Applying Itô's lemma, this implies

$$0 \geq \partial_t v - (\partial_z v) \rho(\hat{x} - \theta) + (\partial_{\hat{x}} v)^\top F \hat{x} + \frac{1}{2} \text{Tr}(\Sigma_t \partial_{\hat{x}\hat{x}} v), \quad (2)$$

with equality on  $\mathfrak{R}$  (when it is not optimal to change  $\theta$ ), where

$$\Sigma_t = K_t R_t K_t^\top = P_t A_t^\top R_t^{-1} A_t P_t^\top. \quad (3)$$

Considering the possibility of changing  $\theta$ , we observe

$$v(t, z, \hat{x}, \theta; \lambda) \geq \sup_{\theta'} v(t, z - \lambda, \hat{x}, \theta'; \lambda) \quad (4)$$

with equality on the complement  $\mathfrak{R}^c$  (when it is optimal to change  $\theta$ ). Combining these inequalities, we obtain the dynamic programming equation

$$0 = \min \left\{ -\partial_t v + (\partial_z v) u(\hat{x} - \theta) - (\partial_{\hat{x}} v)^\top F \hat{x} - \frac{1}{2} \text{Tr}(\Sigma_t \partial_{\hat{x}\hat{x}} v), \right. \\ \left. v(t, z, \hat{x}, \theta; \lambda) - \sup_{\theta'} v(t, z - \lambda, \hat{x}, \theta'; \lambda) \right\}, \quad (5)$$

with terminal value  $v(T, z, \hat{x}, \theta; \lambda) = U(z)$ . The difficulty is that the free boundary for  $\mathfrak{R}$  needs to be determined as part of the solution.

#### 3.1 Asymptotic analysis

When  $\lambda = 0$ , one can use  $\Theta_t = \hat{X}_t$  to achieve  $v(t, z, \hat{x}; 0) \equiv U(z)$ . We expect that the optimal strategy<sup>2</sup> will involve switching whenever  $|\hat{x} - \theta| = O(\lambda^{1/4})$ , resulting in a cost of  $O(\lambda^{1/2})$ . This gives the ansatz

$$v(t, z, \hat{x}, \theta; \lambda) = U(z) - \lambda^{1/2} \phi(t, z) - \lambda \psi(t, z, \hat{x}, \xi) + O(\lambda^{3/2}) \quad (6)$$

where  $\xi := \lambda^{-1/4}(\hat{x} - \theta)$ . See Muhle-Karbe et al. (2017) for further discussion.

<sup>2</sup>This strategy comes from analyzing, over long horizons, how often the boundary of an interval will be hit by a random walk, averaging out the cost paid, then optimizing over the width of the interval chosen. The scaling arguments of Altarovici et al. (2015) and Lo et al. (2004) also apply in our setting, mutatis mutandis. A corresponding analysis in discrete time can be found in the appendix.

Recalling our assumptions on  $\rho$ ,

$$\rho(\hat{x} - \theta) = \rho(\lambda^{1/4}\xi) = \lambda^{1/2}\xi^\top \Gamma \xi + o(\lambda^{1/2}),$$

where

$$\Gamma = \frac{\partial_{xx}\rho(0)}{2}$$

is a positive-definite matrix. We substitute the ansatz (6) into (2), to obtain

$$0 \leq \lambda^{1/2} \left( \partial_t \phi + (\xi^\top \Gamma \xi) U' + \frac{1}{2} \text{Tr}(\Sigma_t \partial_{\xi\xi} \psi) \right) + o(\lambda^{1/2})$$

with equality on  $\mathfrak{K}$ . From (4) we have

$$\begin{aligned} 0 &\leq v(t, z, \hat{x}, \theta; \lambda) - \sup_{\theta'} v(t, z - \lambda, \hat{x}, \theta'; \lambda) \\ &= \lambda U'(z) - \lambda \left( \psi(t, z, \hat{x}, \xi) - \inf_{\xi'} \psi(t, z - \lambda, \hat{x}, \xi') \right) + o(\lambda) \end{aligned} \quad (7)$$

with equality on  $\mathfrak{K}^c$ . For small  $\lambda$ ,

$$\inf_{\xi'} \psi(t, z - \lambda, \hat{x}, \xi') = \psi(t, z - \lambda, \hat{x}, 0) = o(\lambda).$$

This is because  $\psi$  is multiplied by  $\lambda$  in (6) (so its value at any fixed point is irrelevant to first order) and  $\psi$  is smooth. Consequently, (7) simplifies to

$$0 \leq \lambda \left( U'(z) - \psi(t, z, \hat{x}, \xi) \right) + o(\lambda).$$

The leading-order terms for small  $\lambda$  in each region in turn lead to the following approximate version of the dynamic programming equation (5):

$$0 = \min \left\{ \partial_t \phi + (\xi^\top \Gamma \xi) U'(z) + \frac{1}{2} \text{Tr}(\Sigma_t \partial_{\xi\xi} \psi), \quad U'(z) - \psi(t, z, \hat{x}, \xi) \right\}. \quad (8)$$

### 3.2 Exponential Utility

To obtain a closed-form solution to (8), assume that  $U(z) = (1 - e^{-kz})/k$  for some  $k > 0$ . Then  $U'(z) = e^{-kz}$  and with  $\tilde{\phi} = e^{kz}\phi$ ,  $\tilde{\psi} = e^{kz}\psi$ , (8) becomes

$$0 = \min \left\{ \partial_t \tilde{\phi} + \xi^\top \Gamma \xi + \frac{1}{2} \text{Tr}(\Sigma_t \partial_{\xi\xi} \tilde{\psi}), \quad 1 - \tilde{\psi}(t, z, \hat{x}, \xi) \right\}.$$

Following Atkinson and Wilmott (1995), we propose<sup>3</sup> a solution of the form

$$\tilde{\psi}(t, z, \hat{x}, \xi) = -1 + (\xi^\top M \xi - 1)^2$$

with  $\mathfrak{K} = \{\xi : \xi^\top M \xi < 1\}$ , where  $M$  is a (symmetric, positive-definite) matrix to be determined. We have

$$\partial_{\xi\xi} \tilde{\psi} = 4(\xi^\top M \xi - 1)M - 8M\xi\xi^\top M.$$

<sup>3</sup>In one dimension, this is the smallest family of polynomials satisfying our assumptions which are smooth across the boundary.

Therefore,

$$\begin{aligned} 0 &= \partial_t \tilde{\phi} + \xi^\top \Gamma \xi + \frac{1}{2} \text{Tr}(\Sigma_t (4(\xi^\top M \xi - 1)M - 8M \xi \xi^\top M)) \\ &= \partial_t \tilde{\phi} - 2 \text{Tr}(\Sigma_t M) + \xi^\top \left( \Gamma + 2M \text{Tr}(\Sigma_t M) - 4M \Sigma_t M \right) \xi. \end{aligned}$$

This has to hold for all  $\xi$ , so

$$0 = \Gamma + 2M \text{Tr}(\Sigma_t M) - 4M \Sigma_t M \quad (9)$$

which is an algebraic equation for  $M$ . In one dimension, (9) simplifies to  $M = \sqrt{\Gamma/(2\Sigma_t)}$ , and using (3) the approximately optimal no-switching region is

$$\mathfrak{K} = \left\{ |\xi| \leq \left( \frac{\Sigma_t}{\Gamma/2} \right)^{1/4} \right\} = \left\{ |\hat{x} - \theta| \leq \sqrt{P_t} \left( \frac{A_t}{\sqrt{R_t}} \right)^{1/2} \left( \frac{2\lambda}{\Gamma} \right)^{1/4} \right\}.$$

## 4 Interpretation as hypothesis testing

We now explore the connection with hypothesis testing in the case where  $X$  and  $Y$  are scalar processes.

*Example 2.* Suppose we are in the setting of Example 1, so  $X_t$  is constant,  $F = Q = 0$ ,  $A = R = 1$ , and consider testing the hypothesis

$$H_0 : \mu = \Theta_t \quad \text{vs.} \quad H_1 : \mu \neq \Theta_t$$

where  $\mu = X_t$  for  $t \geq 0$ . Recall that

$$P_t = \Sigma_t^{1/2} = (1/P_0 + t)^{-1}$$

is the variance of the hidden process  $X_t$  given the observations  $\mathcal{F}_t$  until time  $t$ , and our asymptotically optimal policy is to switch whenever

$$\frac{|\hat{X}_t - \Theta_t|}{\sqrt{P_t}} > c$$

for  $c = (2\lambda/\Gamma)^{1/4}$ . By choosing the test size of a two-sided test such that  $c$  is the critical value of the usual test statistic, we equivalently switch whenever the standard  $z$ -test rejects  $H_0$ .

In the general scalar setting, the optimal switching region still corresponds to a hypothesis test, but with *variable* test size. Indeed,  $\mathfrak{K}$  then has width proportional to

$$\Sigma_t^{1/4} = \sqrt{P_t} \left( \frac{A_t}{\sqrt{R_t}} \right)^{1/2}. \quad (10)$$

Observe that  $A_t/\sqrt{R_t}$  describes the quality of observations (it is the infinitesimal signal/noise ratio) and hence the volatility of  $\hat{X}$ . Therefore, in periods of low-quality data our agent switches more frequently, or equivalently, uses a test with lower confidence level.

## References

- A. B. Abel, J. C. Eberly, and S. Panageas. Optimal inattention to the stock market with information costs and transaction costs. *Econometrica*, 81(4): 1455–1481, 2013.
- A. Altarovici, J. Muhle-Karbe, and H. M. Soner. Asymptotics for fixed transaction costs. *Finance Stoch.*, 19(2):363–414, 2015.
- F. Alvarez, F. Lippi, and J. Passadore. Are state and time dependent models really different? In M. Eichenbaum and J. Parker, editors, *NBER Macroeconomics Annual 2016*, volume 31, pages 379–457. University of Chicago Press, 2017.
- C. Atkinson and P. Wilmott. Portfolio management with transaction costs: an asymptotic analysis of the Morton and Pliska model. *Math. Finance*, 5(4): 357–367, 1995.
- A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering*. Springer, New York, 2009.
- A. S. Caplin and D. F. Spulber. Menu costs and the neutrality of money. *QJE*, 102(4):703–725, 1987.
- C. Carroll. Macroeconomic expectations of households and professional forecasters. *QJE*, 118(1):269–298, 2003.
- M.-R. Feodoria. *Optimal investment and utility indifference pricing in the presence of small fixed transaction costs*. PhD thesis, Christian-Albrechts-Universität zu Kiel, 2016.
- L. Huang and H. Liu. Rational inattention and portfolio selection. *JF*, 62(4): 1999–2040, 2007.
- R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *J. Basic Eng.*, 83(1):95–108, 1961.
- R. Korn. Portfolio optimisation with strictly positive transaction costs and impulse control. *Finance Stoch.*, 2(2):85–114, 1998.
- A. W. Lo, H. Mamaysky, and J. Wang. Asset prices and trading volume under fixed transactions costs. *JPE*, 112(5):1054–1090, 2004.
- J. Magnani, A. Gorry, and R. Oprea. Time and state dependence in an Ss decision experiment. *AEJ: Macroeconomics*, 8(1):285–310, 2016.
- G. D. Menzies and D. J. Zizzo. Inferential expectations. *B.E.J. Macroeconomics*, 9(1):1–25, 2009.
- J. Muhle-Karbe, H. M. Soner, and M. Reppen. A primer on portfolio choice with small transaction costs. *Ann. Rev. Fin. Econ.*, 9:301–331, 2017.

## 5 Appendix: Illustration in discrete time

In order to give a concrete example of an estimation problem, we consider the task of estimating the parameter  $p \in (0, 1)$  of independent Bernoulli trials  $\{Y_n\}_{n \in \mathbb{N}}$  with values in  $\{0, 1\}$ . We write  $\mathcal{F}_t$  for the information available from the first  $t$  observations, that is  $Y_1, \dots, Y_t$ . Our analysis of this problem will also yield an asymptotic approximation with the same form as (6).

The maximum likelihood estimator (MLE) for the parameter  $p$  is  $\hat{p}_t = \sum_{n=1}^t Y_n/t$ . If  $t$  is large, we apply the central limit theorem, and so have the approximate distribution

$$\hat{p}_t \stackrel{\text{approx.}}{\sim} N\left(p, \frac{p(1-p)}{t}\right).$$

Consider testing

$$H_0 : p = p_0 \quad \text{vs.} \quad H_1 : p \neq p_0.$$

A two-sided hypothesis test with confidence level  $\alpha$  is a rule whereby we maintain  $H_0$  if  $\hat{p}_t$  falls into a confidence interval. This standard ‘belief band of inaction’ is given by

$$p_0 - z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{t}} < \hat{p}_t < p_0 + z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{t}}, \quad (11)$$

where  $z_{\alpha/2}$  is the appropriate quantile of a standard normal distribution. We now demonstrate exactly the same  $1/\sqrt{t}$  scaling effect from a very different perspective, as in the main body of the paper.

Suppose our agent uses an estimator  $\Theta_t$ , based on the sample  $\{Y_i\}_{i=1}^t$ . She incurs two costs:

- A cost  $\lambda > 0$  whenever  $\Theta_t$  changes;
- A cost  $\rho(\Theta_t - \hat{p}_t)$ , based on the deviation of  $\Theta_t$  from the MLE  $\hat{p}_t$ , paid at every time. We assume  $\rho$  is twice differentiable, convex and has a minimum at  $\rho(0) = 0$ .

We ignore the possible wealth effects and risk aversion, and simply aim to minimize the expected future cost. This is equivalent to  $U(z) \equiv z$  (or sending the risk aversion  $k \rightarrow 0$ ) in the setting adopted in the main text.

*Remark 1.* The cost  $\rho$  can be motivated in various ways. One approach is to treat the true probability  $p$  in a Bayesian fashion, and assume our agent faces a running cost  $E[(p - \Theta_t)^2 | \mathcal{F}_t]$ , that is, a cost depending on the distance of their estimate from the true (unknown) value. This is a classic example of a Bayesian loss function, and is suggested by the agent facing additional risk (as measured by the conditional variance) whenever their estimate deviates from  $p$ .

In this case, the MLE satisfies  $\hat{p}_t = E[p | \mathcal{F}_t]$ , and we can compute

$$E[(p - \Theta_t)^2 | \mathcal{F}_t] = E[(p - \hat{p}_t)^2 | \mathcal{F}_t] + (\hat{p}_t - \Theta_t)^2.$$



As the agent has no control over the term  $E[(p - \hat{p}_t)^2 | \mathcal{F}_t]$ , the effective cost is given by  $(\hat{p}_t - \Theta_t)^2$ , which is of the form considered.

The cost  $\lambda$  can be motivated, for example by assuming that the choice of  $\Theta$  is an input to a more complex decision setting, which will need to be recalibrated whenever  $\Theta$  changes.

Related results for models with transaction costs suggest that the optimal policy is for the agent not to act until the tracking error  $\hat{p}_t - \Theta_t$  leaves some interval. To a first-order approximation, which we consider more formally below, the interval is of the form  $(-b_t, b_t)$ , for some  $b_t$  to be determined. When the difference between the MLE and the agent's approximation exceeds this threshold, the optimal strategy is to set  $\Theta_t = \hat{p}_t$  (this is because  $\hat{p}_t$  is an unbiased estimate of  $p$  and the adjustment cost does not depend on the size of the adjustment).

To find  $b_t$ , we will work in an asymptotic regime, where we consider small costs  $\lambda \rightarrow 0$  and fix a large terminal time  $T \rightarrow \infty$ . In particular, we assume that as  $T \rightarrow \infty$ , there is a time  $t_* \rightarrow \infty$  such that the long-run costs on the interval  $[t_*, T]$  form the principal part of the realized cost, and costs on the interval  $[0, t_*]$  can be ignored.

We first consider the behaviour of the tracking error between two sequential switching times  $t_1 < t_2$ , where  $t_* \leq t_1$ . (Note that  $t_* \rightarrow \infty$  as  $T \rightarrow \infty$ , so  $t_1 \rightarrow \infty$  by assumption.) We assume that

$$b_{t_1} \rightarrow 0, \quad t_1 b_{t_1} \rightarrow \infty, \quad \frac{b_t}{b_{t_1}} \rightarrow 1 \quad \text{and} \quad \frac{t b_t}{t_1 b_{t_1}} \rightarrow 1 \quad \text{for all } t \in [t_1, t_2] \quad (12)$$

as  $T \rightarrow \infty, \lambda \rightarrow 0$ . We shall see that these assumptions are consistent with the optimal  $b_t$  we construct.

We can write

$$\hat{p}_t - \hat{p}_{t_1} = \frac{1}{t} \sum_{i=t_1+1}^t (Y_i - \hat{p}_{t_1}) \approx \frac{1}{t_1} \sum_{i=t_1+1}^t (Y_i - \hat{p}_{t_1}), \quad (13)$$

where the approximation is justified whenever  $t_1^{-1} - t_2^{-1}$  is small, which is justified as  $t_1 \rightarrow \infty$ . Assuming  $t_1$  is large<sup>4</sup>, we know that

$$P(Y_i = 1) = p \approx \hat{p}_t \approx \hat{p}_{t_1},$$

so the tracking error is approximately the sum of a sequence of mean-zero iid random variables, and is well modelled as a random walk, with each step having up-probability  $\hat{p}_{t_1}$ .

For notational simplicity, we write  $\hat{\sigma}_t^2 = \hat{p}_t(1 - \hat{p}_t)$ , which is the estimated variance of our observations. Note that  $\hat{\sigma}_t^2 \rightarrow p(1 - p)$  as  $t \rightarrow \infty$ , in particular  $\hat{\sigma}_t \approx \hat{\sigma}_{t_*}$  when  $T$  (and hence  $t_*$ ) is large.

Using the approximation of tracking error as a random walk, we choose  $b_t$  to minimize expected costs. We have to trade off between our running cost and

<sup>4</sup>This calculation is for a classical/frequentist approach. With a Bayesian approach we would instead have  $P(Y_i = 1 | \mathcal{F}_{i-1}) = \hat{p}_i \approx \hat{p}_{t_1}$ , giving the same asymptotic approximation.

the cost of switching. For a time  $s$ , we try and evaluate the expected cost at time  $t$ , given our barrier strategy  $b_t$ . We first compute the running cost term.

Write  $C_t^\rho(b_t) = \rho(\hat{p}_t - \Theta_t)$  when  $\Theta_t$  is determined using a boundary  $b_t$ . From our assumptions on  $\rho$ , provided the tracking error is not too large (which will happen whenever  $b_t$  is small), we can approximate with Taylor's theorem  $\rho(x) \approx \Gamma x^2$  for some constant  $\Gamma > 0$ .

As  $T$  is large, our agent will be active over a long horizon, so it is the long-run average value of this cost which is important. As  $b_t$  will change through time, it is natural to rescale our random walk, so we look for the asymptotic stationary distribution of

$$\xi_t = (\hat{p}_t - \hat{p}_{t_1})/b_{t_1}.$$

This is given approximately by the 'triangular' density

$$g(\xi) = \begin{cases} 1 + \xi & \text{if } -1 < \xi \leq 0, \\ 1 - \xi & \text{if } 0 < \xi \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

as can be seen by the observations that:

- $\xi_t$  jumps to zero whenever it hits  $\pm b_t/b_{t_1} \rightarrow \pm 1$ , so  $g(-1) = g(1) = 0$ .
- The density integrates to unity.
- From considering the possible paths of  $\xi$ : Except at  $x = 0$ , the only way for  $\xi$  to reach  $x$  is from being previously at either  $x - \frac{1-\hat{p}_t}{t_1 b_{t_1}}$  and observing  $Y = 1$ , or at  $x + \frac{\hat{p}_{t_1}}{t_1 b_{t_1}}$  and then observing  $Y = 0$ . The probability of observing  $Y = 1$  is  $\hat{p}_{t_1}$ , so we have the stationary Chapman–Kolmogorov equation

$$g(x) = \hat{p}_{t_1} g\left(x - \frac{1-\hat{p}_{t_1}}{t_1 b_{t_1}}\right) + (1-\hat{p}_{t_1}) g\left(x + \frac{\hat{p}_{t_1}}{t_1 b_{t_1}}\right) \quad \text{for } x \neq 0.$$

Rearranging and writing  $h = 1/(t_1 b_{t_1})$ , we obtain

$$0 = \hat{p}_{t_1} \frac{g(x - (1-\hat{p}_{t_1})h) - g(x)}{h^2} + (1-\hat{p}_{t_1}) \frac{g(x + \hat{p}_{t_1}h) - g(x)}{h^2} \quad \text{for } x \neq 0.$$

Assuming  $g$  is twice differentiable for  $x \neq 0$  and taking  $h \rightarrow 0$  (or equivalently  $t_1 b_{t_1} \rightarrow \infty$ ), this gives the differential equation

$$0 = g'(x) \left( -\hat{p}_{t_1}(1-\hat{p}_{t_1}) + (1-\hat{p}_{t_1})\hat{p}_{t_1} \right) h^{-1} \\ + \frac{g''(x)}{2} \left( \hat{p}_{t_1}(1-\hat{p}_{t_1})^2 + (1-\hat{p}_{t_1})\hat{p}_{t_1}^2 \right) + o(1) \quad \text{for } x \neq 0,$$

or equivalently  $g''(x) = 0$  for  $x \neq 0$ . Hence, when  $t_1 b_{t_1}$  is large (which is assumed when  $T \rightarrow \infty$ ,  $\lambda \rightarrow 0$ ), we obtain a piecewise linear function for  $g$ , and hence the triangular density.

The density  $g(\cdot)$  has variance  $1/6$  and substituting,

$$E[C_t^\rho(b_t)|\mathcal{F}_{t_*}] \approx \Gamma \frac{b_t^2}{6}$$

for  $t_* \ll t$ .

We now seek to understand the expected switching cost,  $E[C_t^\lambda(b_t)|\mathcal{F}_{t_*}]$ , where  $C_t^\lambda(b_t) = \lambda$  if  $|\xi_t| \geq 1$  and zero otherwise. Using our rescaled random walk  $\xi$ , we need to find the probability of  $\xi$  hitting  $\pm b_t/b_{t_1} \approx \pm 1$  at a time  $t \gg t_*$ . As  $\xi$  is approximately a random walk restarted at zero, this is approximately  $1/E[\tau|\mathcal{F}_{t_1}]$ , where  $\tau \approx t_2 - t_1$  is the time taken to hit  $\pm 1$  from zero by  $\xi$ .

To calculate  $E[\tau|\mathcal{F}_{t_1}]$ , first observe that  $\xi_t^2 - ct$  is approximately a martingale for  $c = \hat{\sigma}_{t_1}^2/(t_1 b_{t_1})^2 \approx \hat{\sigma}_{t_*}^2/(t b_t)^2$  and  $\xi$  will eventually hit  $\pm 1$ . By the optional stopping theorem,

$$\xi_{t_1}^2 - ct_1 \approx E[\xi_{t_2}^2 - ct_2|\mathcal{F}_{t_1}] = E[1 - ct_2|\mathcal{F}_{t_1}].$$

Rearrangement, and the fact  $\xi_{t_1} = 0$ , yields  $1 \approx cE[t_2 - t_1|\mathcal{F}_{t_1}] = cE[\tau|\mathcal{F}_{t_1}]$ .

For  $t_* \ll t$ , with  $t_1 < t \leq t_2$ , the probability of hitting the barrier at  $t$ , thus incurring cost  $\lambda$ , is  $1/E[\tau|\mathcal{F}_{t_*}]$ , so

$$E[C_t^\lambda(b_t)|\mathcal{F}_{t_*}] \approx \lambda \left( \frac{\hat{\sigma}_{t_*}}{t b_t} \right)^2 \quad \text{for } t_* \ll t.$$

We can now minimize our expectations of long-run future costs. At the leading order, using our approximations, the costs are given by

$$E[C_t^\lambda(b_t)|\mathcal{F}_{t_*}] + E[C_t^\rho(b_t)|\mathcal{F}_{t_*}] \approx \lambda \left( \frac{\hat{\sigma}_{t_*}}{t b_t} \right)^2 + \Gamma \frac{b_t^2}{6}.$$

Minimizing this expression in a pointwise manner gives

$$b_t = \left( \frac{6\lambda}{\Gamma} \right)^{1/4} \sqrt{\frac{\hat{\sigma}_{t_*}}{t}} = \chi \frac{\hat{\sigma}_{t_*}}{\sqrt{t}} \quad \text{where } \chi = \left( \frac{6\lambda}{\Gamma} \right)^{1/4} \hat{\sigma}_{t_*}^{-1/2}.$$

Therefore, we obtain a bandwidth  $b_t$  for the belief band of inaction that is proportional to  $1/\sqrt{t}$ , giving the same behaviour as a confidence interval.

As in the continuous time case, we also observe that the width of our band of inaction is  $\chi \hat{\sigma}_{t_*}/\sqrt{t}$ , where  $\hat{\sigma}_{t_*}/\sqrt{t}$  is the standard deviation of the classical estimator. The coefficient  $\chi$  is determined by the square root of the signal/noise ratio  $\sigma_{t_*}^{-1/2}$ . (To see that this corresponds to the signal/noise ratio, observe that  $\text{Var}(\hat{p}_t - \hat{p}_{t-1}) \approx \hat{\sigma}_{t_*}/t$ , while the variance of  $\hat{p}_t$  is  $\hat{\sigma}_{t_*}^2/t$ . Their ratio,  $\hat{\sigma}_{t_*}^{-1}$ , then corresponds to the term  $A_t/\sqrt{R_t}$  from the Kalman–Bucy dynamics (1), which appears in the corresponding band of inaction (10).)

As in the continuous case, we observe that our band of inaction corresponds with a confidence interval, and hence to a two-sided hypothesis test, with test size dependent on the signal/noise ratio. Using this  $b$ , we also have

$$E[C_t^\lambda(b_t)|\mathcal{F}_{t_*}] + E[C_t^\rho(b_t)|\mathcal{F}_{t_*}] \approx \frac{\hat{\sigma}_{t_*}}{t} \sqrt{\frac{2\lambda\Gamma}{3}}$$

which agrees with the  $\lambda^{1/2}$  scaling of our asymptotic approximation in continuous time (6).

Finally, we can check the consistency of our choice of  $b_t$  with our asymptotic assumption (12). Clearly, we have  $b_t \rightarrow 0$  and  $tb_t \rightarrow \infty$  as  $t \rightarrow \infty$ . We also know that the period between consecutive switches  $\tau$  has expectation  $E[\tau|\mathcal{F}_{t_*}] \approx (tb_t/\hat{\sigma}_{t_*})^2 = O(\lambda^{1/2}t)$ . As  $\lambda \rightarrow 0$ , this shows that, for consecutive switching times  $t_1, t_2$ , for  $t \in [t_1, t_2]$ ,

$$\frac{b_t}{b_{t_1}} \approx \frac{(t_1 + O(\lambda^{1/2}t_1))^{1/2}}{t_1^{1/2}} \rightarrow 1 \quad \text{and} \quad \frac{tb_t}{t_1 b_{t_1}} \approx \frac{(t_1 + O(\lambda^{1/2}t_1))^{3/2}}{t_1^{3/2}} \rightarrow 1.$$

Our choice of  $b_t$  is thus consistent with our asymptotic assumption (12) as  $T \rightarrow \infty, \lambda \rightarrow 0$ .