Optimal Discrete Adjustments for Short Production Runs

Scott A. Vander Wiel

AT&T Bell Laboratories
Murray Hill, New Jersey 07974

January 19, 1993

ABSTRACT

Diameter measurements on successive metal hubs from a machining operation are modeled using a random walk with observation error and linear drift corresponding to tool wear. After producing and measuring a hub the depth of the cutting tool can be adjusted an integer number of tenths (0.0001 inches.) How should the tool be adjusted?

This paper studies a version of the problem omitting tool wear. The objective is to minimize run costs proportional to the sum of squared diameter deviations from a target plus fixed charges for tool adjustments. The optimal strategy makes no adjustment if an estimate of the process mean is near target. Otherwise, an adjustment is made to return the estimated mean as near to target as possible within the adjustment resolution.

The region where no adjustments are made widens near the end of the production run where adjustments have only short term impact. The region converges as the number of remaining periods increases. Plots of expected run costs show that the extra cost of discreteness is small at high resolution but is substantial when the adjustment grid is coarse.

Keywords: feedback control, minimum cost, fixed adjustment cost

1 INTRODUCTION

In a particular machining operation producing metal hubs, adjustments to the depth of a cutting tool are made in multiples of 1 tenth (0.0001 inch.) This compares to a specification tolerance on the hub diameter of ± 5 tenths and typical diameter ranges of less than 2 tenths in samples of, say, 5 consecutive parts. On this basis it is natural to question whether the discreteness of tool adjustments substantially interferes with the goal of machining parts with diameters near a target value.

Of course, the answer depends on many factors such as the meaning of "substantially interfere" and the stochastic structure (if any) of consecutive measurements. For the machining operation it is reasonable to model consecutive part diameters using a random walk with observation error and linear drift (corresponding to tool wear.) If performance is measured by the sum of a fixed adjustment cost and a variable (for example, squared deviation) off target cost, it is possible to derive an optimal tool adjustment strategy for a simpler model omitting tool wear. The optimal strategy adjusts only if the current estimate of the location of the random walk deviates from the target diameter by more than a certain amount. An optimal nonzero adjustment is to compensate (as nearly as possible) for the deviation of the current estimate from its target level. Numerical calculations of optimal expected costs help quantify the intuitive result that coarser adjustment resolutions give rise to increasingly higher expected production costs.

Section 2 of this paper describes the hub machining operation. The random walk model with observation error and linear tool wear is shown to provide a reasonable fit to a set of successive diameter measurements. Section 3 discusses the optimal adjustment strategy for the fitted model ignoring tool wear. Plots are provided showing deadband limits (within which no adjustments are made) and optimal expected costs for various levels of adjustment resolution. Section 4 is an analytical derivation of the optimal discrete adjustment policy using dynamic programming. Section 5 studies convergence of the deadband limits and expected cost functions as the number of periods remaining in a run increases. A summary and review of some relevant literature is given in Section 6.

2 A MACHINING APPLICATION

This section describes an actual machining step in the production of metal hubs. Data from a study of the process is presented and a portion of this data is modeled as a random walk with observation error and linear drift.

2.1 Description

A horizontal bar lathe at a manufacturing site produces metal hubs (see the simplified diagram in Figure 1). Finished hubs can be loaded into an automatic gaging device to measure several key part dimensions including the outside diameter. The nominal diameter is 1.501 inches with a specification tolerance of ± 5 tenths. The gaging process is known to contribute moderately to the observed variation of diameter measurements. Figure 2 is a plot of outside diameter measurements obtained on 150 consecutive hubs machined during a study of the machine tool. Measurements were taken on every hub produced.

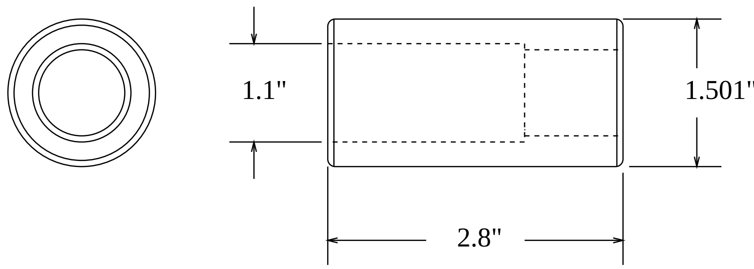

Figure 1: Diagram of metal hub

During the 7 1/2 hour study period, one adjustment (of -8 tenths) was made to the cutting tool position and this took place after the 75th hub was turned. The values plotted in Figure 2 are corrected to show the diameters which "would" have resulted had no offset been made. It is significant that the smallest adjustment which can be made on this tool is 1 tenth and that this is of the same order of magnitude as typical diameter ranges from small samples of consecutive parts. Most ranges lie between 0.5 and 1.5 tenths for samples of 5. Thus, it is sensible to explicitly consider adjustment discreteness in this application.

Several aspects of the machining operation are immediately clear from the plot in Figure 2. During production of the first 126 hubs it is evident that the diameters tend to increase roughly linearly. This can be attributed to progressive wear of the cutting tool. In fact, because of wear, the tool is usually (though not during the study) replaced with a refurbished one after approximately every 75 hubs. Secondly, after hub 126 the diameters jump dramatically and show a downward trend for about 20 hubs. This behavior is explained by a 20 minute break taken by the operator. During the break the lathe was turned off and the hydraulic system cooled from its "steady state" temperature. Cooler hydraulics reduced the force holding the cutting tool in position and thus rendered the tool effectively further from the workpiece. The effect appears to have decreased as the system warmed to its normal operating temperature over the final 24 parts in the study.

2.2 State Space Model

Transient effects such as hydraulic warm up are not considered in this paper. Hence, the final 24 hub diameter measurements following the operator's break are not included in the following analysis.

The top two panels of Figure 3 show the sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF) of residuals from a simple linear regression fit using ordinary least squares. The hashed lines are approximate 2 standard error bands appropriate if the residuals are uncorrelated. Clearly there is predictive information in the measurements beyond that of a simple linear trend.

A simple state space time series model proposed for this data set is that of a random walk with drift and observation error. In the following specification the units of measurement are tenths.

$$\theta_t \quad = \quad d + \theta_{t-1} + u_{t-1} + \nu_t, \tag{1}$$

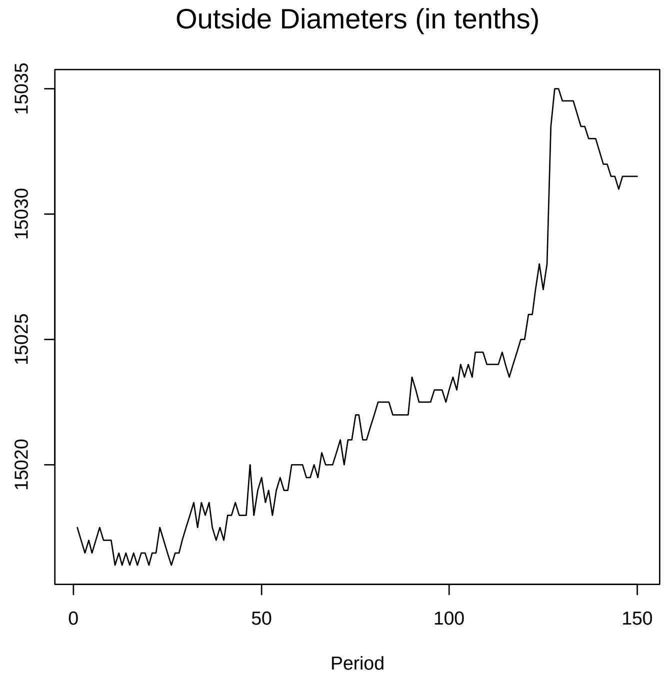

Figure 2: Outside diameters measurements from 150 consecutive hubs

$$y_t = \theta_t + \epsilon_t \tag{2}$$

where

$$\begin{aligned} \theta_t &= \text{unobserved "true" diameter deviation from 15010 (target)}, \\ d &= \text{tool wear per hub (drift)}, \\ u_{t-1} &= \text{offset made just after hub } t-1, \\ \nu_1, \nu_2, \ldots &\sim \text{iid N}(0, \sigma_\nu^2), \\ y_t &= \text{measured diameter deviation from 15010}, \\ \epsilon_1, \epsilon_2 &\sim \text{iid N}(0, \sigma_\epsilon^2) \end{aligned}$$

and the ν_t sequence is independent of the ϵ_t sequence. For the purpose of estimation, the model is viewed conditionally on y_1=7.5. This is a useful initialization since (1) and (2) alone do not define a joint distribution for a sequence of observations $(y_1, \ldots, y_n)$. Different choices for initializing the model have very little effect on parameter estimates.

Maximum likelihood estimates and approximate 95% confidence intervals (in parentheses) based on the first 126 diameter measurements are

$$\hat{\sigma}_\epsilon = 3.75, \ \ (3.02, 4.65), \ \ \hat{\sigma}_\nu = 3.10, \ \ (2.27, 4.24), \ \ \hat{d} = 0.84, \ \ (0.29, 1.38).$$

The confidence intervals are based on the observed information matrix parameterized with $\ln \sigma_\epsilon$, $\ln \sigma_\nu$ and d. Harvey (1989) discusses estimation using the Kalman filter. The lower four panels of Figure 3 present diagnostics for the state space model using the standardized residuals. Sample ACF and PACF functions are shown along with a scatter plot and a normal probability plot. The model fits reasonably well with respect to these diagnostics and for further illustration it is assumed to be correct. It is particularly noteworthy that the parameter

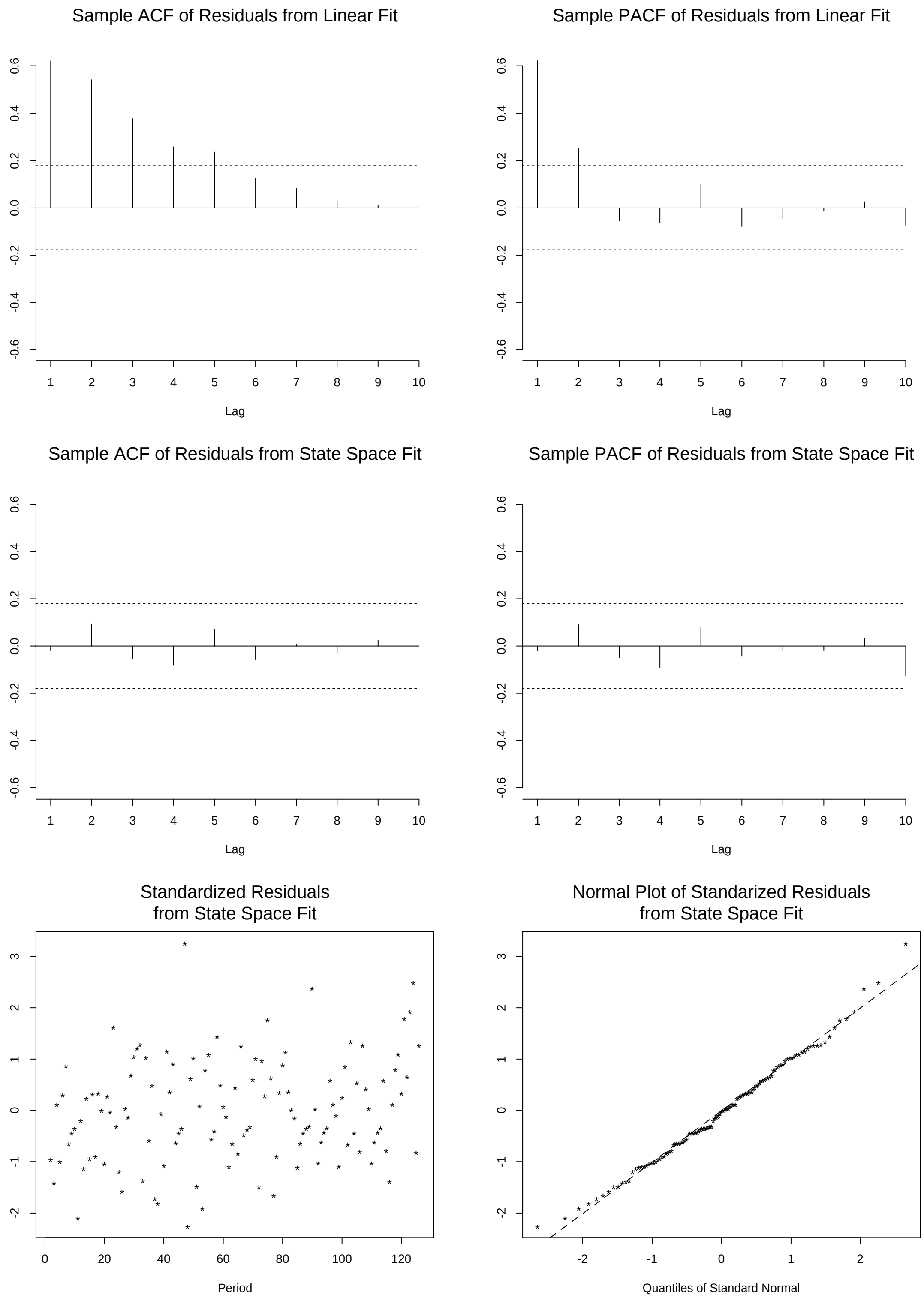

Figure 3: Some diagnostics on residuals from simple linear regression and statespace models

estimates are all on the same order of magnitude as the smallest possible adjustment of 1 tenth. This suggests that adjustment discreteness may not be negligible.

2.3 Control Objective

Since it is desirable to operate the bar lathe so as to machine hubs with diameters close to the nominal value, it seems reasonable to build an objective function including costs monotone in $|\theta_t|$. In addition, it should be recognized that fewer and smaller adjustments are generally more desirable and hence it is reasonable to include costs monotone in $|u_{t-1}|$. For the metal machining operation we consider modeling the total expected cost over a run of n hubs in proportion to

$$E\left\{\sum_{t=1}^{n}[\theta_t^2 + K_a\delta(u_{t-1})]\right\} \tag{3}$$

where

$$\begin{aligned}\delta(u) &= 0, \quad u = 0\\ &= 1, \quad u \neq 0.\end{aligned}$$

Thus, run costs in period t are modeled proportional to θ_t^2 and a fixed cost $K_a > 0$ for a nonzero adjustment. Control engineers would more typically use a factor u_{t-1}^2 in place of $\delta(u_{t-1})$. However, (3) better models a situation in which minor adjustments are just as time consuming and disruptive as are major adjustments. This is true, for example, when an adjustment requires a fixed amount of process down time and labor as is the case for some preventive maintenance procedures.

3 OPTIMAL DISCRETE CONTROL

This section draws on results derived in Sections 4 and 5 to present the optimal solution to the bar lathe control problem for the case of no tool wear (ie, $d = 0$). Of course tool wear is actually nonnegligible and hence the solution pertains to a problem simpler than the real one. Jensen (1989) includes linear tool wear but assumes adjustments are continuous. The case containing both tool wear and discrete adjustments is difficult to handle analytically and is a topic for further research.

3.1 Deadband Scheme

Under the convenient initialization $\theta_0 \sim \mathrm{N}(\hat{\theta}_0, q_\infty)$ where q_∞ is given by (6), the system (1) and (2) with $d = 0$ is (by Corollary 1) adjusted optimally with respect to the objective function (3) by setting

$$\begin{aligned}u_{t-1} &= 0, && \text{if } |\hat{\theta}_{t-1}| \le k_{n-t+1}\\ &= <-\hat{\theta}_{t-1}>, && \text{otherwise}\end{aligned} \tag{4}$$

where $<\theta>$ is the integer nearest θ and

$$\hat{\theta}_{t-1} = E\{\theta_{t-1}|y_1, \ldots, y_{t-1}\}.$$

The constants k_t depend only on K_a and σ_ν^2 (not σ_ϵ^2). The conditional expectation $\hat{\theta}_{t-1}$ can be computed recursively using (7) below.

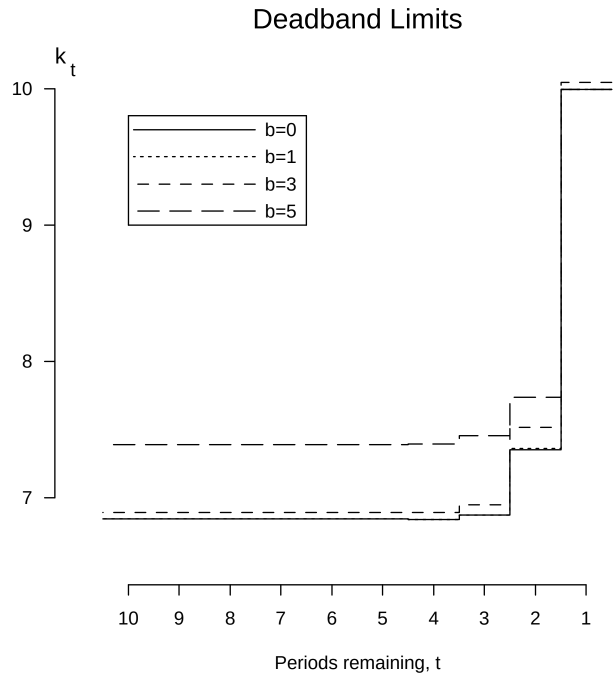

Figure 4: Optimal deadband limits for various adjustment precisions

The optimal scheme makes no adjustment if the machine is perceived to be operating near its target level (ie, if $\hat{\theta}_{t-1}$ lies near 0). However, if $\hat{\theta}_{t-1}$ is sufficiently far from 0, an adjustment is made to offset the perceived misadjustment as nearly as possible within the constraints of the the discrete manipulator variable u_{t-1}. The constants k_t are indexed in reverse order (that is k_n is used in period $t = 1$) because they are obtained through a recursion which first generates the constant k_1 appropriate for period n. See Theorem 3 and Corollary 1. The optimal adjustment strategy is very similar to that of Crowder (1991) for the case of continuous adjustments. In his strategy the nonzero adjustment is simply $-\hat{\theta}_{t-1}$ and the constants k_t take on different values. He noted that the optimal control scheme is very similar to a traditional exponentially weighted moving average chart used in statistical process monitoring.

If the smallest absolute tool adjustment were b tenths (as opposed to 1 tenth) then by measuring in units of b tenths the problem can be transformed to the unit resolution case discussed above. Figure 4 is a plot of the deadband limits k_t against t for various levels of adjustment resolution and n=10. Figures 4 through 6 use $\sigma_\nu = 3.10$ from the bar lathe application along with $K_a = 100$ and they are constructed so as to apply for any value of σ_ϵ. The case $b = 0$ corresponds to Crowder's (1991) continuous adjustment problem. The figure shows generally that increased granularity in the adjustment variable leads to larger values of k_t. That is, for more granular adjustments, one should perceive a larger deviation from target before making an adjustment to the tool position. Interestingly the differences are greater for longer horizons (that is, near the beginning of the run.) Also, as Crowder explained, the limits fan out near the end of the run so that the estimated mean must be further from target before making a costly adjustment that can have only short term impact. Furthermore, the deadbands appear to converge to the left as the horizon length increases. See Theorem 6.

3.2 Optimal Risk Functions

Using $\theta_0 \sim \mathrm{N}(\hat{\theta}_0, q_\infty)$ is called steady state initialization. From standard Kalman filtering theory (eg, Meinhold and Singpurwalla, 1983) or Bayes theorem, one obtains conditional distributions

$$\begin{aligned} \theta_{t-1}|(y_1,\ldots,y_{t-1}) &\sim \mathrm{N}(\hat{\theta}_{t-1}, q_\infty), \\ \theta_t|(y_1,\ldots,y_{t-1}) &\sim \mathrm{N}(\hat{\theta}_{t-1}+u_{t-1}, q_\infty+\sigma_\nu^2), \quad \text{and} \\ \hat{\theta}_t|(y_1,\ldots,y_{t-1}) &\sim \mathrm{N}(\hat{\theta}_{t-1}+u_{t-1}, \sigma_\nu^2) \end{aligned} \tag{5}$$

where

$$q_\infty = \sigma_\nu^2 \left[\left(\frac{1}{4} + \frac{\sigma_\epsilon^2}{\sigma_\nu^2} \right)^{1/2} - \frac{1}{2} \right], \tag{6}$$

$$\hat{\theta}_t = \rho_\infty(\hat{\theta}_{t-1}+u_{t-1}) + (1-\rho_\infty)y_t, \quad \text{and} \tag{7}$$

$$\rho_\infty = \frac{\sigma_\epsilon^2}{q_\infty + \sigma_\nu^2 + \sigma_\epsilon^2}\,.$$

That the conditional variances do not depend on t is due to the initialization using q_∞. More generally, if $\theta_0 \sim \mathrm{N}(\hat{\theta}_0, q_0)$, the posterior variances approach their steady state values exponentially fast.

An important sequence of functions in understanding the optimal adjustment strategy is the sequence of optimal risk functions $R_n(\hat{\theta}_0)$ giving the expected cost incurred in running the process optimally for n periods with an initial mean $\hat{\theta}_0$. That is,

$$R_n(\hat{\theta}_0) = \min_{u_0,\ldots,u_{n-1}} E\left\{ \sum_{t=1}^{n} \theta_t^2 + K_a\delta(u_{t-1}) \right\}.$$

The minimization is over the functions u_{t-1} mapping $(y_1,\ldots,y_{t-1},u_0,\ldots,u_{t-2})$ to the integers and the expectation is with respect to the (joint) distribution of $(y_1,\ldots,y_n,\theta_1,\ldots,\theta_n)$.

Suppose the $n-1$ period problem is solved and $R_{n-1}(\hat{\theta}_0)$ is known. The optimal adjustment u_0 for the first period of an n period problem represents a tradeoff in the expected costs of two alternatives. On the one hand, if no adjustment is made ($u_0 = 0$) the expected run cost using (3) and (5) is

$$\begin{aligned} R_n(\hat{\theta}_0|u_0=0) &= E\{\theta_1^2 + R_{n-1}(\hat{\theta}_1)\} \\ &= \hat{\theta}_0^2 + q_\infty + \sigma_\nu^2 + \int R_{n-1}(\hat{\theta}_1)h(\hat{\theta}_1;\hat{\theta}_0)\,d\hat{\theta}_1 \end{aligned}$$

where $h(x;\mu)$ is the $\mathrm{N}(\mu,\sigma_\nu^2)$ density. On the other hand if the adjustment $u_0 = \langle -\hat{\theta}_0 \rangle$ is made the expected run cost is

$$\begin{aligned} R_n(\hat{\theta}_0|u_0=\langle -\hat{\theta}_0 \rangle) &= K_a + E\{\theta_1^2 + R_{n-1}(\hat{\theta}_1)\} \\ &= K_a + r^2(\hat{\theta}_0) + q_\infty + \sigma_\nu^2 + \int R_{n-1}(\hat{\theta}_1)h(\hat{\theta}_1; r(\hat{\theta}_0))\,d\hat{\theta}_1 \end{aligned}$$

where

$$r(\hat{\theta}_0) = \hat{\theta}_0 + \langle -\hat{\theta}_0 \rangle$$

is the (signed) distance to the integer nearest $\hat{\theta}_0$. The n period optimal risk $R_n(\hat{\theta}_0)$ is the minimum of these expected costs

$$R_n(\hat{\theta}_0) = \min\{R_n(\hat{\theta}_0|u_0 = 0), R_n(\hat{\theta}_0|u_0 = <-\hat{\theta}_0>)\}$$

and an optimal u_0 is correspondingly 0 or $<-\hat{\theta}_0>$.

The three panels on the left of Figure 5 are plots of $R_n(\hat{\theta}_0) - R_n(0)$ for $n \in \{1,2,3\}$ and $b \in \{0,1,3)\}$. They illustrate the shapes of the optimal risk functions. not depend on The three center panels show more detail of the functions in the regions of transition. The three panels on the right are plots of $R_n(0) - np$ (where $p = q_\infty + \sigma_\nu^2$) against n and show how the level of the optimal risk increases as the horizon lengthens. Though the sequence $R_n(0)$ depends on σ_ϵ^2, the sequence $R_n(0) - np$ does not (see Lemma 6) in the appendix.

Consider, for example, $R_2(\hat{\theta}_0) - R_2(0)$ for $b = 1$ plotted in the left center panel. In the neighborhood of $\hat{\theta}_0 = 0$, $R_2(\hat{\theta}_0) = R_2(\hat{\theta}_0|u_0 = 0)$ and the shape is dominated by the term $\hat{\theta}_0^2$. The interpretation of this is that if the process mean is perceived to be near its target value, no adjustment is made and future costs are nearly quadratic in $\hat{\theta}_0$.

Far enough from $\hat{\theta}_0 = 0$ the function $R_2(\hat{\theta}_0)$ is wave-like — periodic in $\hat{\theta}_0$ and roughly quadratic on the intervals $[i - 1/2, i + 1/2]$ with local minima at the integers i. This is because far from $\hat{\theta}_0 = 0$, $R_2(\hat{\theta}_0) = R_2(\hat{\theta}_0|u_0 = <-\hat{\theta}_0>)$ and the shape is dominated by the term $r^2(\hat{\theta}_0)$ which has a quadratic wave shape. The wave can be understood intuitively by realizing that if an integer adjustment is to be made, then starting with an estimated mean $\hat{\theta}_0$ is no different (in terms of optimal expected cost) from starting at $\hat{\theta}_0 + i$ for any integer i. Furthermore, if $\hat{\theta}_0$ happens to be an integer, then the adjustment $u_0 = <-\hat{\theta}_0> = -\hat{\theta}_0$ will return the estimated mean to its target value and this is the best possible position. However, if $\hat{\theta}_0$ is merely in the neighborhood of an integer, the adjustment $u_0 = <-\hat{\theta}_0>$ can only return the estimated mean to the neighborhood of zero. Hence, the optimal expected cost should be roughly quadratic in the neighborhood of integers (far from zero) with local minima at those integers. Furthermore, for large $|i|$ (and any n), $R_n(i) - R_n(0) = K_a$ which is 100 in the example.

The two points of transition between the near-quadratic and wave portions of $R_n(\hat{\theta}_0)$ are the deadband limits $\pm k_n$. They are the break-even points between making no adjustment and making an adjustment $u_0 = <-\hat{\theta}_0>$. By Theorem 7 the functions $R_n(\hat{\theta}_0) - R_n(0)$ converge uniformly as $n \to \infty$ and this accounts for the convergence of the deadband limits k_t (Theorem 6) observed in Figure 4.

In Figure 5 an obvious difference in the shapes of the risk functions for $b = 1$ and $b = 3$ is that the period of the wave portions changes from 1 to 3. With continuous adjustments ($b = 0$) the "wave" is constant since then the estimated mean can be returned exactly to zero regardless of $\hat{\theta}_0$. It is also apparent that the peaks of the $b = 3$ curves are higher than those of the $b = 1$ curves. This is because with $b = 3$ the worst possible $\hat{\theta}_0$ (a large odd multiple of 1.5) leaves the estimated mean (after adjustment) three times as far from zero as the corresponding worst case with $b = 1$.

3.3 Effect of Horizon Length

There are also interesting features in the plots on the right side of Figure 5. The horizontal axis indicates the number of periods n in the production run so it is expected that the total cost would increase to the right as more periods are added. Although not perceptible, the

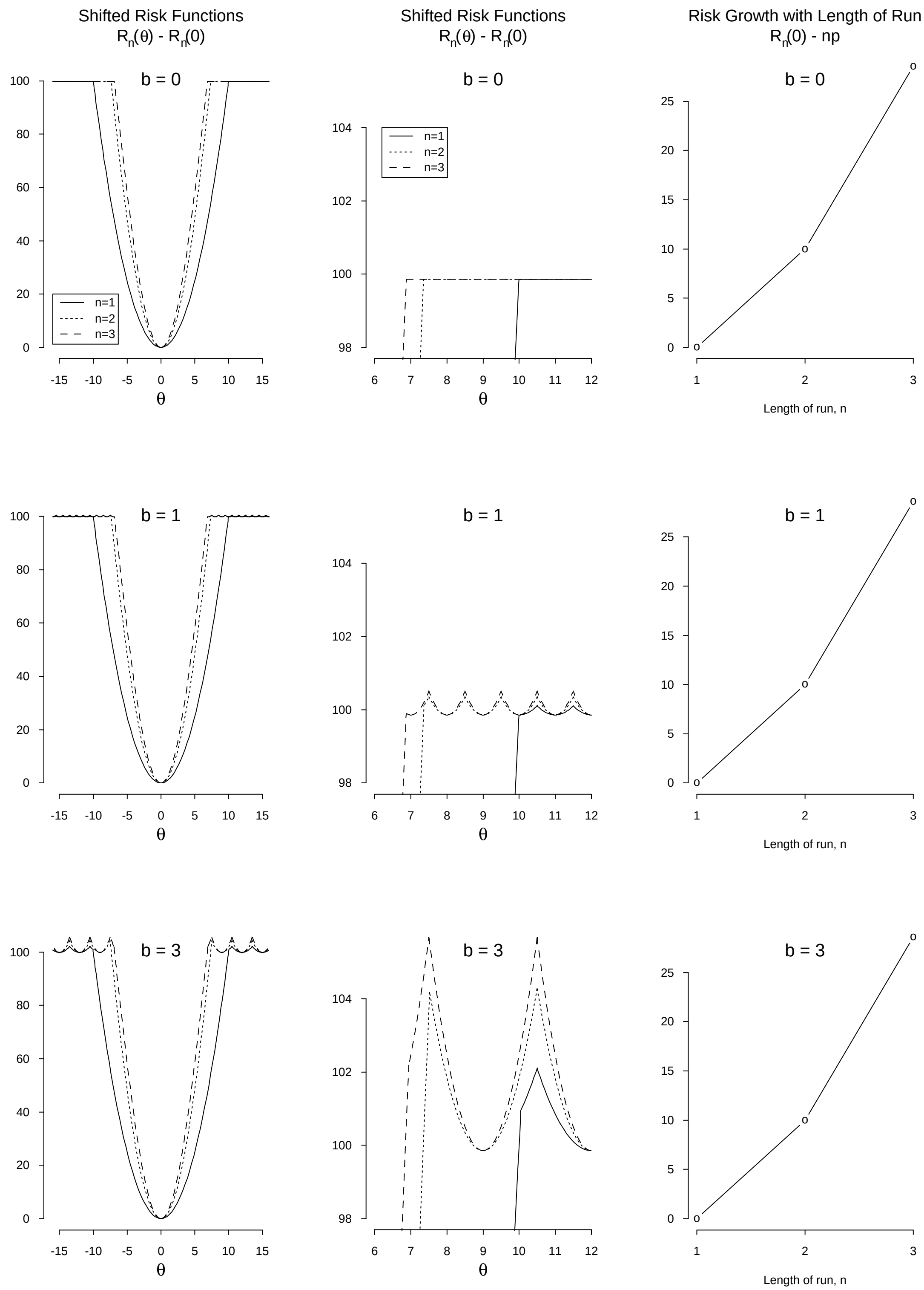

Figure 5: Optimal risk functions for various adjustment precisions

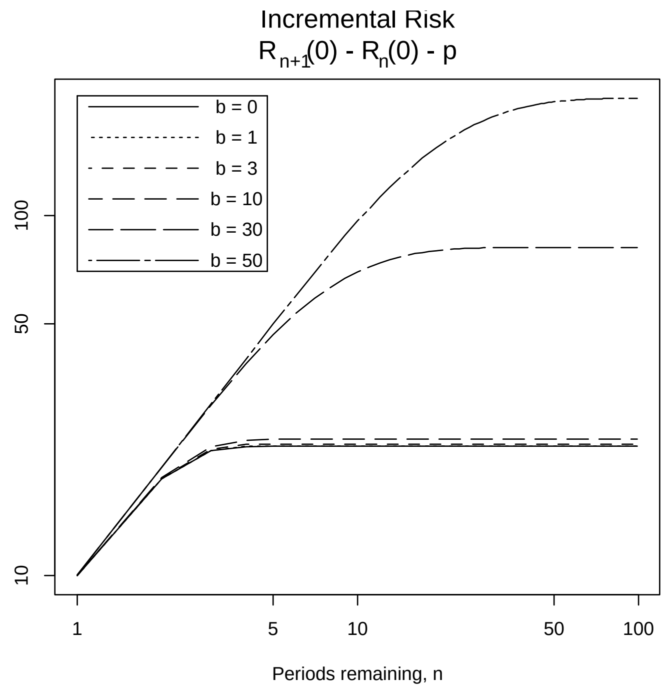

Figure 6: Incremental risk functions for various adjustment precisions

curves rise faster as b increases. The convex shapes of the curves indicate that the marginal cost of adding an extra period to the run increases with the length of the run. This makes intuitive sense because longer horizon problems will likely have a larger fraction of periods in which costly adjustments are made since these adjustments can have longer term impact. However, one would expect this increase in incremental expected cost to eventually level off. This would correspond to the curves approaching linear asymptotes as n increases.

Plotted against n in Figure 6 is the incremental risk defined as $R_n(0) - R_{n-1}(0) - p$. These curves give the slopes of those in the right panels of Figure 5. Both axes are on log scales. For illustrative purposes the figure includes some relatively coarse values of the adjustment precision b. The figure shows that relatively fine discrete adjustments increase expected costs only slightly over the continuous adjustment case. But coarse adjustments increase expected costs substantially for each period added to the production run. Also, the incremental risk appears to converge as the horizon becomes longer (as conjectured in the preceding paragraph.) The convergence is slower for coarser adjustment resolution (larger b) and very slow for very large b. Another interesting feature is that the incremental risk sequences are nearly identical and increase nearly linearly for problems with short horizons or very coarse adjustments. This can be understood by realizing that in these cases the optimal policy will often make no adjustment during a short production run begun on target. Thus costs will accrue almost solely from squared deviations of the mean from the target value. For an unadjusted random walk θ_t, $E\{\theta_t^2\}$ is linearly increasing in t.

3.4 General Comments Regarding Discrete Adjustments

Whereas the actual machining application described in Section 2 has deterministic drift due to tool wear, this paper studies a no drift version of the discrete control problem. Nevertheless, the solution provides some justification for one's intuition regarding optimal discrete control of the drift model (1) and (2) with respect to the objective function (3). In particular, Figure 6 has shown (for $d = 0$) that the extra cost of discreteness is small if the adjustments are made with high resolution. In this case the system can be adequately approximated as having a continuous adjustment variable. It seems reasonable that this would also hold for the drift version of the problem at least for small $|d|$. If incremental expected costs for the version with drift are similar to those shown in Figure 6 the cost of discreteness at the level of tenths is not severe for the hub machining application.

Crowder (1991) studied a *continuous* adjustment version of the machine control problem without drift. The present strategy (4) is very close to a rounded version of his solution. The only difference in addition to rounding the optimal adjustments is that the deadband limits are somewhat different. Jensen (1989) studied a continuous adjustment version of the machine control problem with drift. In this case the optimal adjustment strategy still has the deadband form. However, the deadband region is shifted in the direction of $-d$ and is thus no longer symmetric about zero. Furthermore, the optimal nonzero adjustment becomes

$$u_{t-1} = -\hat{\theta}_{t-1} - d - z^*_{n-t+1}$$

where z^*_{n-t+1} has the sign of d. The optimal adjustment consists of (i) a term $(-\hat{\theta}_{t-1})$ to correct for the perceived misadjustment, (ii) a term $(-d)$ to compensate for the drift to occur in period t and (iii) an overcompensation term $(-z^*_{n-t+1})$ which anticipates the drift to occur in future periods which transpire before the next adjustment. In the absence of a solution to the full problem with discrete adjustments *and* drift, it seems that a reasonable strategy would be to round Jensen's drift adjustment solution, possibly using slightly wider deadband limits. It would be interesting to compare this strategy to the (as yet unstudied) optimal one.

Finally, it is interesting that setting the fixed adjustment cost K_a to zero reduces the control objective to minimizing the sum of mean squared errors of θ_t. If the adjustment variable is continuous, the well known optimal strategy is $u_{t-1} = -\hat{\theta}_{t-1}$. The solution in the discrete adjustment case is simply the rounded version $u_{t-1} = <-\hat{\theta}_{t-1}>$. A common objective function in engineering control literature replaces the fixed adjustment cost $K_a\delta(u_{t-1})$ in (3) with a term $K_a u^2_{t-1}$. This is know as the linear quadratic Gaussian problem and penalizes the magnitude (instead of the number) of adjustments. It is not generally true that rounding the optimal continuous adjustment solution under this objective produces optimal discrete adjustments. However, it seems reasonable that the increase in expected costs would not be great, especially for high resolution adjustments.

4 DERIVATION

This section derives the optimal adjustment policy (4) under a cost model slightly more general than (3). The steps are as follows. First, Theorem 1 gives important properties of the optimal risk functions $R_n(\theta)$. These properties allow an explicit recursive equation for generating the sequence $R_n(\theta)$. Theorem 2 gives a bound on $|\hat{\theta}_0|$ not depending on n, beyond which it is optimal to make the nonzero adjustment $u_0 = <-\hat{\theta}_0>$.

Theorem 3 provides a three term decomposition of $R_n(\theta)$. The sequence of remainder functions in the decomposition is the negative part of a sequence $g_n(\theta)$ which satisfies a simple recursive equation. The sign of $g_n(\hat{\theta}_0)$ determines whether the optimal adjustment u_0 is zero or $<-\hat{\theta}_0>$. If each $g_n(\theta)$ crosses zero only once in its positive domain, the optimal adjustment strategy has the deadband form (4).

4.1 Problem Statement

Let process measurements y_t evolve according to (1) and (2) with $d = 0$ and initial condition $\theta_0 \sim \mathrm{N}(\hat{\theta}_0, q_\infty)$. The discrete adjustment problem is to determine adjustments $u_{t-1} \in \mathcal{Z}$ ($\mathcal{Z}$ is the set of integers) based on $(y_1, \ldots, y_{t-1})$ and $(u_0, \ldots, u_{t-2})$ so as to minimize the expected run cost given by

$$L(n) = E\left\{\sum_{t=1}^{n}[C_0(\theta_t) + K_a\delta(u_{t-1})]\right\} \tag{8}$$

where $K_a > 0$ and the off target cost function $C_0(\theta)$ is assumed to satisfy

1. $C_0(\theta)$ is nonnegative and symmetric about $\theta = 0$,
2. $C_0(\theta) \leq C_0(\theta + 1)$ for every $\theta > -1/2$,
3. $\lim_{|\theta|\to\infty} C_0(\theta) = \infty$, and
4. $\int C_0(x)\phi(x; \theta, q_\infty + \sigma_\nu^2)\, dx$ is finite for all θ,

where $\phi(x; \mu, \sigma^2)$ is the $\mathrm{N}(\mu, \sigma^2)$ density. The cost model (3) used in the lathe application corresponds to $C_0(\theta) = \theta^2$. However, (8) allows other off target costs such as $C_0(\theta) = |a\theta|$ or $C_0(\theta) = a_0 \ln(1 + |a\theta|)$ for $a_0 > 0$.

An apparently more general problem would allow adjustments to be integer multiples of a constant b. However, by scaling to units of b, the problem can be expressed in terms of integer adjustments. Further standardization is possible by noting that minimizing $L(n)$ is no different from minimizing a positive multiple of $L(n)$. Thus, K_a can be taken as 1 without loss of generality. Deadband limits and optimal expected costs for the general problem can be obtained using

$$\begin{aligned} k_t\,(b, K_a, C_0(x), \sigma_\nu) &= bk_t\,(1, 1, C_0(bx)/K_a, \sigma_\nu/b) \quad \text{and} \\ R_n\,(\theta; b, K_a, C_0(\theta), \sigma_\nu, \sigma_\epsilon) &= K_a R_n\,(\theta/b; 1, 1, C_0(bx)/K_a, \sigma_\nu/b, \sigma_\epsilon/b) \end{aligned}$$

where the dependence of k_t and $R_n(\theta)$ on b, K_a, C_0, σ_ν^2 and σ_ϵ^2 has been made explicit and the conditions on $C_0(\theta)$ are modified appropriately. The results below assume the problem has been scaled to have integer adjustments.

4.2 Optimal Risk Functions

This section develops a recursive equation for the sequence of optimal risk functions $R_n(\theta)$ and inductively establishes important properties of these functions.

The posterior distributions given in (5) imply that optimal adjustments u_{t-1} depend on the process history only through the posterior mean $\hat{\theta}_{t-1} = E\{\theta_{t-1}|y_1, \ldots, y_{t-1}\}$. Define

$R_0(\theta) = 0$. Using the dynamic programming principle of optimality, $R_n(\theta)$ can be written in terms of $R_{n-1}(\theta)$ $(n \geq 1)$ as follows.

$$
\begin{aligned}
R_n(\theta) &= \min_{u \in \mathcal{Z}} \left\{ \int [C_0(x) + K_a \delta(u)] \phi(x; \theta + u, \sigma_\nu^2 + q_\infty)\, dx \right. \\
&\quad \left. + \int R_{n-1}(x) \phi(x; \theta + u, \sigma_\nu^2)\, dx \right\} \\
&= \min \left\{ \min_{u \in \mathcal{Z}} \left[C(\theta + u) + K_a + \int R_{n-1}(x) h(x; \theta + u)\, dx \right]; \right. \\
&\quad \left. C(\theta) + \int R_{n-1}(x) h(x; \theta)\, dx \right\}
\end{aligned}
\tag{9}
$$

where

$$
C(\theta) = \int C_0(x) \phi(x; \theta, \sigma_\nu^2 + q_\infty)\, dx, \tag{10}
$$

and henceforth

$$
h(x; \theta) = \phi(x; \theta, \sigma_\nu^2)
$$

is the $\mathrm{N}(\theta, \sigma_\nu^2)$ density. Some properties of $C(\theta)$ that follow from Lemma 2 in the appendix and the assumed properties of $C_0(\theta)$ are

1. $C(\theta)$ is nonnegative and symmetric about $\theta = 0$,
2. $C(\theta) < C(\theta + 1)$ for every $\theta > -1/2$,
3. $\lim_{|\theta| \to \infty} C(\theta) = \infty$, and
4. $C(\theta)$ is continuous in θ.

The following theorem gives properties of the optimal risk functions and makes the minimization over $u \in \mathcal{Z}$ in (9) explicit.

Theorem 1 *For each $n \geq 1$,*

1. *$R_{n-1}(\theta)$ is nonnegative, symmetric about $\theta = 0$ and such that $R_{n-1}(\theta) \leq R_{n-1}(\theta + 1)$ for every $\theta > -1/2$.*
2. *An integer u minimizing $C(\theta + u) + K_a + \int R_{n-1}(x) h(x; \theta + u)\, dx$ is $u = {<}{-\theta}{>}$.*
3. *$R_n(\theta)$ is given by*

$$
\begin{aligned}
R_n(\theta) &= \min \left\{ C(r(\theta)) + K_a + \int R_{n-1}(x) h(x; r(\theta))\, dx; \right. \\
&\quad \left. C(\theta) + \int R_{n-1}(x) h(x; \theta)\, dx \right\}.
\end{aligned}
\tag{11}
$$

4. *It is optimal in an n-period problem is to take u_0 equal to ${<}{-\hat{\theta}_0}{>}$ or 0 according to whether $R_n(\hat{\theta}_0)$ is equal to the first or second minimand in (11).*

Proof The proof of 1 and 2 is by induction. Consider the case $n = 1$. Then 1 holds since $R_0(\theta) = 0$. In 2 the quantity to be minimized over integers u is

$$C(\theta + u) + K_a.$$

This is symmetric in $\mu = \theta + u$ about the point $\mu = 0$ and for $\mu > -1/2$ satisfies

$$C(\mu) + K_a \leq C(\mu + 1) + K_a.$$

Thus, Lemma 3 in the appendix implies

$$\min_{u \in \mathcal{Z}} \{C(\theta + u) + K_a\} = C(r(\theta)) + K_a.$$

Since the integer $u = <-\theta>$ gives $\theta + u = r(\theta)$, this u is a minimizer.

Now suppose 1 and 2 hold for an $(n-1)$ period problem. It will be shown that they must also hold for an n-period. By Lemma 2 in the appendix, the function

$$G(\mu) = \int R_{n-1}(x) h(x; \mu)\, dx$$

is nonnegative, symmetric and such that $G(\mu) \leq G(\mu + 1)$ for every $\mu > 1/2$. These properties also hold for $C(\mu) + K_a + G(\mu)$. Hence, Lemma 3 in the appendix implies

$$C(\theta + u) + K_a + \int R_{n-1}(x) h(x; \theta + u)\, dx$$

is minimized by the integer $u = <-\theta>$ so that $\theta + u = r(\theta)$. Therefore

$$\min_{u \in \mathcal{Z}} \left\{ C(\theta + u) + K_a + \int R_{n-1}(x) h(x; \theta + u)\, dx \right\} = C(r(\theta)) + K_a + G(r(\theta)). \tag{12}$$

Since C, K_a and G are nonnegative and symmetric functions then so is their sum. Furthermore, since $r(\theta) = r(\theta + 1)$, the minimum (12) is trivially nondecreasing on any $\{\theta, \theta + 1\}$. These properties of (12), combined with equation (9), Lemma 4 in the appendix and the fact that

$$C(\theta) + \int R_{n-1}(x) h(x; \theta)\, dx$$

is nonnegative, symmetric about $\theta = 0$ and nondecreasing on $\{\theta, \theta + 1\}$ for every $\theta > -1/2$, establish that $R_n(\theta)$ is nonnegative, symmetric about $\theta = 0$ and nondecreasing on $\{\theta, \theta + 1\}$ for every $\theta > -1/2$.

Parts 3 and 4 follow from parts 1 and 2 by equation (9). □

The next theorem gives a bound on the magnitude of $\hat{\theta}_0$ outside of which an initial nonzero adjustment is optimal regardless of the horizon length n.

Theorem 2 *In an n period problem if $|\hat{\theta}_0|$ exceeds*

$$k = \inf\{\theta > 0 : C(x) - C(r(x)) \geq K_a,\ \forall x \geq \theta\},$$

it is optimal to take $u_0 = <-\hat{\theta}_0>$.

Proof By Theorem 1 (parts 3 and 4) it must be shown that for $\theta > k$,

$$C(\theta) - C(r(\theta)) - K_a + \int R_{n-1}(x)[h(x;\theta) - h(x;r(\theta))]\,dx \geq 0. \tag{13}$$

Clearly the sum of the first three terms is nonnegative for every $\theta > k$. By Theorem 1 (part 1) and Lemma 2 in the appendix

$$\int R_{n-1}(x)h(x;\theta)\,dx$$

is nonnegative, symmetric and nondecreasing in θ on $\{\theta, \theta+1\}$ for every $\theta > -1/2$. Also,

$$\int R_{n-1}(x)h(x;r(\theta))\,dx$$

has the same value for θ and $\theta + 1$. Hence, for any $\theta > -1/2$

$$\int R_{n-1}(x)[h(x;\theta) - h(x;r(\theta))]\,dx \tag{14}$$

is nondecreasing on $\{r(\theta), r(\theta)+1, \ldots, \theta\}$. But since (14) is a symmetric function of θ and is zero at $r(\theta)$, it is nonnegative for any $\theta \in \mathcal{R}$. Therefore (13) holds for every $\theta > k$. $\Box$

The next theorem (proved in the appendix) provides a three term decomposition of $R_n(\theta)$ into (i) a constant multiple of n, (ii) a function periodic in θ, and (iii) a remainder. The remainder functions are the negative parts of a sequence $g_n(\theta)$ which satisfies a simple recursive equation. Also, the sign of $g_n(\hat{\theta}_0)$ determines whether the optimal adjustment u_0 is zero or $<-\hat{\theta}_0>$.

Theorem 3 *Let $g_0(\theta) = v_0(r(\theta)) = 0$ and for $n \geq 1$ define recursively*

$$g_n(\theta) = [C(\theta) - C(r(\theta)) - K_a] + \int g_{n-1}^-(x)[h(x;\theta) - h(x;r(\theta))]\,dx \tag{15}$$

and

$$v_n(r(\theta)) = [C(r(\theta)) - C(0)] + \int [v_{n-1}(r(x)) + g_{n-1}^-(x)]h(x;r(\theta))\,dx$$

where $g_n^-(\theta) = \min\{0, g_n(\theta)\}$.

Then for $n \geq 1$,

1. $R_n(\theta) = n[C(0) + K_a] + v_n(r(\theta)) + g_n^-(\theta)$,
2. *$g_n(\theta)$ is continuous, symmetric about $\theta = 0$ and strictly increasing on $\{\theta, \theta+1\}$, $\forall \theta > -1/2$,*
3. *an optimal first period adjustment in an n period problem is to take u_0 to be zero or $<-\hat{\theta}_0>$ according as $g_n(\hat{\theta}_0)$ is negative or nonnegative.*

The following straightforward corollary shows that if each $g_n(\theta)$ crosses zero only once, then optimal adjustments are nonzero only when $|\hat{\theta}_{t-1}|$ exceeds a deadband limit.

Corollary 1 *Suppose the sets $\{|\theta| : g_n(|\theta|) = 0\}$ are nonempty (possibly degenerate) intervals, say, $[k_{L,n}, k_{U,n}]$, and choose $k_n \in [k_{L,n}, k_{U,n}]$. Then an optimal adjustment for period t ($t = 1, \ldots, n$) of an n period problem is*

$$\begin{aligned} u_{t-1} &= 0, && \text{if } |\hat{\theta}_{t-1}| \leq k_{n-t+1} \\ &= <-\hat{\theta}_{t-1}>, && \text{otherwise.} \end{aligned}$$

Furthermore, the integral in (15) need only cover the range finite range $[-k_{n-1}, k_{n-1}]$ contained in $[-k, k]$ where k is defined in Theorem 2.

The zero crossing condition is easily checked numerically and if it holds the limits k_n can be obtained by straightforward numerical integration using, for example, a quadrature technique. The condition has been found to hold throughout extensive computations over a wide range of parameter values for the quadratic loss $[C_0(\theta) = \theta^2]$ version of (8).

5 LONG HORIZON CONVERGENCE

Section 4 studied the discrete adjustment problem with a focus on finite production runs. This section considers whether the adjustment policy converges as the horizon length n increases. In particular, convergence of the shifted optimal risk functions $R_n(\theta) - R_n(0)$ is studied in two stages. Write

$$R_n(\theta) - R_n(0) = f_n(\theta) + w_n(r(\theta))$$

where

$$\begin{aligned} f_n(\theta) &= R_n(\theta) - R_n(r(\theta)), \quad \text{and} \\ w_n(r(\theta)) &= R_n(r(\theta)) - R_n(0). \end{aligned}$$

Theorem 4 below shows that $f_n(\theta)$ converges uniformly to a continuous function $f(\theta)$. Convergence of $f_n(\theta)$ implies (Theorem 5) convergence of the functions $g_n(\theta)$ defined in Theorem 3 and hence convergence of the optimal deadband limits k_n (of Corollary 1) when they exist.

Convergence of $f_n(\theta)$ is thus quite useful. However, it may also be of interest to know when $R_n(\theta) - R_n(0)$ converges. Theorem 7 gives a sufficient condition for uniform convergence of $w_n(r(\theta))$ and hence that of $R_n(\theta) - R_n(0)$. It is conceivable that the condition always holds; it has been verified numerically for $\sigma_\nu > 0.19$.

5.1 Recursion for $f_n(\theta)$

The recursion (11) for $R_n(\theta)$ may be written as follows.

$$\begin{aligned} R_n(\theta) &= \min\{C(r(\theta)) + K_a + E\{R_{n-1}(r(\theta) + Z)\}; \\ &\qquad C(\theta) + E\{R_{n-1}(\theta + Z)\}\}, \quad n = 1, 2, \ldots, \end{aligned} \tag{16}$$

where the random variable Z is distributed as $\mathrm{N}(0, \sigma_\nu^2)$. Since

$$R_n(r(\theta)) = C(\theta) + E\{R_{n-1}(r(\theta) + Z)\}$$

it follows that

$$f_n(\theta) = \min\{K_a; C_1(\theta) + E\{R_{n-1}(\theta + Z) - R_{n-1}(r(\theta) + Z)\}\} \tag{17}$$

where

$$C_1(\theta) = C(\theta) - C(r(\theta)).$$

Since $C(\theta)$ is continuous, nonnegative, symmetric about zero and has $\lim_{\theta\to\infty} C(\theta) = \infty$, $C_1(\theta)$ also has these properties. Furthermore, $C_1(\theta)$ is zero on the interval $[-1/2, 1/2]$. The following two properties of the remainder function $r(x) = x + <-x>$ are used in Section 5 without further reference: (i) $r(r(x)) = r(x)$, and (ii) $r(r(x) + y) = r(x + y)$.

Using property (ii) and the equality $R_n(\theta) = f_n(\theta) + R_n(r(\theta))$ one may write

$$R_n(\theta + Z) - R_n(r(\theta) + Z) = f_n(\theta + Z) - f_n(r(\theta) + Z).$$

Substituting this into (17) gives the recursion

$$f_n(\theta) = \min\{K_a; C_1(\theta) + E\{f_{n-1}(\theta + Z) - f_{n-1}(r(\theta) + Z)\}\}. \tag{18}$$

5.2 Convergence of Deadband Limits

This subsection proves uniform convergence of $f_n(\theta)$ (Theorem 4) and of $g_n(\theta)$ (Theorem 5) and convergence of the deadband limits k_n (Theorem 6).

Bather's (1963) Lemma 2 can be modified to prove

Lemma 1 *The sequence f_n is uniformly bounded. Furthermore, there is a large enough number ξ such that $f_n(\theta) = K_a$ whenever $|\theta| \geq \xi$.*

The lemma is used to prove the following.

Theorem 4 *The sequence $f_n(\theta)$ converges uniformly to a continuous function $f(\theta)$ which satisfies*

$$f(\theta) = \min\{K_a; C_1(\theta) + E\{f(\theta + Z) - f(r(\theta) + Z)\}\}.$$

Proof The argument is a nontrivial extension of the argument leading to Theorem 1 in Bather (1963). The difference $f_{n+1}(\theta) - f_n(\theta)$ is bounded as follows. If $f_{n+1}(\theta) > f_n(\theta)$ then by (18) $K_a \geq f_{n+1}(\theta) > f_n(\theta)$; that is,

$$f_n(\theta) = C_1(\theta) + E\{f_{n-1}(\theta + Z) - f_{n-1}(r(\theta) + Z)\}.$$

Hence from (18),

$$0 < f_{n+1}(\theta) - f_n(\theta) \leq E\{f_n(\theta + Z) - f_n(r(\theta) + Z) - f_{n-1}(\theta + Z) + f_{n-1}(r(\theta) + Z)\}. \tag{19}$$

On the other hand, if $f_{n+1}(\theta) < f_n(\theta)$ then by (18) $K_a \geq f_n(\theta) > f_{n+1}(\theta)$; that is,

$$f_{n+1}(\theta) = C_1(\theta) + E\{f_n(\theta + Z) - f_n(r(\theta) + Z)\}.$$

Hence,

$$0 > f_{n+1}(\theta) - f_n(\theta) \geq E\{f_n(\theta + Z) - f_n(r(\theta) + Z) - f_{n-1}(\theta + Z) + f_{n-1}(r(\theta) + Z)\}. \tag{20}$$

Next define

$$\begin{aligned} F_n(\theta,\omega,\eta) &= [f_{n+1}(r(\theta)) + \omega + \eta) - f_n(r(\theta)) + \omega + \eta)] \\ &\quad -[f_{n+1}(r(\omega)) + \theta + \eta) - f_n(r(\omega)) + \theta + \eta)] \end{aligned}$$

and for fixed θ, ω, and η denote the two bracketed expressions by Q and S respectively. Define

$$M_n = \sup_{\theta,\omega,\eta} |F_n(\theta,\omega,\eta)|.$$

By Lemma 1, each M_n is finite. It will be shown that $\sum M_n < \infty$.

Suppose, for example, that $F_n(\theta,\omega,\eta) > 0$ implying that $Q > S$. Consider the three possible relationships among Q, S and zero.

Case 1: If $Q > S \geq 0$ then by (19)

$$\begin{aligned} F_n(\theta,\omega,\eta) \leq Q &\leq E\{[f_n(r(\theta)) + \omega + \eta + Z) - f_{n-1}(r(\theta)) + \omega + \eta + Z)] \\ &\quad -[f_n(r(\theta) + \omega + \eta) + Z) - f_n(r(\theta) + \omega + \eta) + Z)]\} \\ &= \int F_{n-1}(\theta, \theta + \omega + \eta, z - \theta) h(z;0)\, dz. \end{aligned}$$

Case 2: If $Q > 0 > S$ then by (19) and (20)

$$\begin{aligned} F_n(\theta,\omega,\eta) \leq Q - S &\leq E\{[f_n(r(\theta)) + \omega + \eta + Z) - f_{n-1}(r(\theta)) + \omega + \eta + Z)] \\ &\quad -[f_n(r(\theta) + \omega + \eta) + Z) - f_n(r(\theta) + \omega + \eta) + Z)]\} \\ &\quad -[f_n(r(\omega)) + \theta + \eta + Z) - f_{n-1}(r(\omega)) + \theta + \eta + Z)] \\ &\quad +[f_n(r(\theta) + \omega + \eta) + Z) - f_n(r(\theta) + \omega + \eta) + Z)]\} \\ &= \int F_{n-1}(\theta,\omega,\eta + z) h(z;0)\, dz. \end{aligned}$$

Case 3: If $0 \geq Q > S$ then

$$\begin{aligned} F_n(\theta,\omega,\eta) \leq -S &\leq E\{-[f_n(r(\omega)) + \theta + \eta + Z) - f_{n-1}(r(\omega)) + \theta + \eta + Z)] \\ &\quad +[f_n(r(\theta) + \omega + \eta) + Z) - f_n(r(\theta) + \omega + \eta) + Z)]\} \\ &= \int F_{n-1}(\theta + \omega + \eta, \theta, z - \theta) h(z;0)\, dz. \end{aligned}$$

On the other hand if $F_n(\theta,\omega,\eta) < 0$ similar reasoning shows

Case 4: If $S < Q \leq 0$ then $-F_n(\theta,\omega,\eta) \leq \int F_{n-1}(\theta+\omega+\eta, \theta, z - \theta) h(z;0)\, dz$.

Case 5: If $S < 0 < Q$ then $-F_n(\theta,\omega,\eta) \leq \int F_{n-1}(\omega, \theta, \eta + z) h(z;0)\, dz$.

Case 6: If $0 \leq S < Q$ then $-F_n(\theta,\omega,\eta) \leq \int F_{n-1}(\omega, \theta + \omega + \eta, z - \omega) h(z;0)\, dz$.

By Lemma 1, if neither $r(\theta) + \omega + \eta$ nor $r(\omega) + \theta + \eta$ lies in $(-\xi, \xi)$ then $F_n(\theta,\omega,\eta) = 0$. On the other hand, if at least one of these quantities lies in $(-\xi,\xi)$ then in each of the six cases the integrand of the bound for $|F_n(\theta,\omega,\eta)|$ vanishes either for all $z < -(2\xi + 1)$ or for all $z > 2\xi + 1$. For the remaining values of z the integrand is bounded by $M_{n-1}h(z;0)$. Define $\rho \in (0,1)$ by

$$\rho = 1 - \min\left\{\int_{-\infty}^{-(2\xi+1)} h(z;0)\,dz; \int_{2\xi+1}^{\infty} h(z;0)\,dz\right\}.$$

We have shown for $n = 1, 2, \ldots$,

$$M_n = \sup_{\theta,\omega,\eta} |F_n(\theta,\omega,\eta)| \le \rho M_{n-1}.$$

This implies

$$M_n \le \rho^n M_0$$

and hence $\sum M_n < \infty$. By definition, $f_0(\theta) = 0 = f_n(r(\theta))$. Hence,

$$f_n(\theta) = \sum_{m=0}^{n-1} f_{m+1}(\theta) - f_m(\theta) = \sum_{m=0}^{n-1} F_m(0,\theta,0)$$

which is the partial sum of a series which converges absolutely and uniformly in θ. Thus define

$$f(\theta) = \lim_{n\to\infty} f_n(\theta)$$

and since each $f_n(\theta)$ is continuous and the convergence is uniform, $f(\theta)$ is also continuous.

It remains only to show that $f(\theta)$ solves the functional equation. Given any $\epsilon > 0$ pick $n = n(\epsilon)$ so large that $|f_m(\theta) - f(\theta)| < \epsilon$ for every $m \ge n$ and for every θ. Then

$$\begin{aligned}
|f(\theta) \;&-\; \min\{K_a, C_1(\theta) + E\{f(\theta+Z) - f(r(\theta)+Z)\}\}| \\
&\le\; |f(\theta) - f_{n+1}(\theta)| \\
&\quad +|\min\{K_a, C_1(\theta) + E\{f_n(\theta+Z) - f_n(r(\theta)+Z)\}\} \\
&\quad -\min\{K_a, C_1(\theta) + E\{f(\theta+Z) - f(r(\theta)+Z)\}\}| \\
&\le\; |f(\theta) - f_{n+1}(\theta)| \\
&\quad +|E\{f_n(\theta+Z) - f(\theta+Z) - f_n(r(\theta)+Z) + f(r(\theta)+Z)\}| \\
&\le\; 3\epsilon.
\end{aligned}$$

□

Theorem 5 *The functions $g_n(\theta)$ defined in Theorem 3 converge uniformly to a function $g(\theta)$ which satisfies*

$$g(\theta) \;= [C(\theta) - C(r(\theta)) - K_a] + \int g^-(x)[h(x;\theta) - h(x;r(\theta))]\,dx.$$

Proof From Theorem 3 (part 1),

$$R_n(\theta) = n[C(0) + K_a] + v_n(r(\theta)) + g_n^-(\theta)$$

so

$$f_n(\theta) = R_n(\theta) - R_n(r(\theta)) = g_n^-(\theta) - g_n^-(r(\theta)) = g_n^-(\theta) - K_a. \tag{21}$$

By Theorem 4, $f_n(\theta)$ converges uniformly to, say, $f(\theta)$. Hence, $g_n^-(\theta)$ converges uniformly to $\gamma(\theta) \equiv f(\theta) + K_a$. Define

$$g(\theta) = [C(\theta) - C(r(\theta)) - K_a] + \int \gamma(\theta)[h(x;\theta) - h(x;r(\theta))]\, dx.$$

Then

$$\begin{aligned} |g_n(\theta) - g(\theta)| &= \left| \int [g_{n-1}^-(x) - \gamma(x)][h(x;\theta) - h(x;r(\theta))]\, dx \right| \\ &\leq \sup_x |g_{n-1}^-(x) - \gamma(x)| \int |h(x;\theta) - h(x;r(\theta))|\, dx \\ &\leq 2 \sup_x |g_{n-1}^-(x) - \gamma(x)| \\ &\to 0. \end{aligned}$$

That is, $g(\theta)$ is the uniform limit of $g_n(\theta)$. Hence $\gamma(x) = g^-(x)$. □

Theorem 6 *Suppose for each n the set $\{|\theta| : g_n(|\theta|) = 0\}$ is a nonempty interval, say, $[k_{L,n}, k_{U,n}]$. Suppose further that $\{|\theta| : g(|\theta|) = 0\}$ is a singleton, say, k^* where $g(\theta) = \lim_{n\to\infty} g_n(\theta)$. Then $[k_{L,i}, k_{U,i}] \to k^*$ and*

$$g(\theta) = [C(\theta) - C(r(\theta)) - K_a] + \int_{k^*}^{k^*} g(x)[h(x;\theta) - h(x;r(\theta))]\, dx.$$

The proof is straightforward using properties of $g_n(\theta)$ from Theorem 3 (part 2).

5.3 Convergence of Risk Functions

Theorem 4 provides uniform convergence of $R_n(\theta) - R_n(r(\theta))$. The interpretation of this is that the *shapes* of the risk functions $R_n(\theta)$ stabilize for large n uniformly on any set of shifted integers $\{z + r : z \in \mathcal{Z}\}$, $r \in [-1/2, 1/2]$ and the convergence is uniform over all such sets. Under an additional condition, the following theorem guarantees the uniform convergence of $R_n(\theta) - R_n(0)$ and yields the interpretation that the shapes of $R_n(\theta)$ stabilize uniformly over the whole real line.

Theorem 7 *Suppose for some positive integer m*

$$\max_{|s|\leq 1/2} I^m(s) < 1$$

where

$$\begin{aligned} I(s) &= \int_{-1/2}^{1/2} |Q(x,s)|\, dx, \\ I^n(s) &= \int_{-1/2}^{1/2} I^{n-1}(x)|Q(x,s)|\, dx, \end{aligned}$$

for $n = 2, 3, \ldots$, *and where*

$$Q(x,s) = \sum_{j=-\infty}^{\infty} h(x+j;s) - h(x+j;0).$$

Then the functions $w_n(s) \equiv R_n(s) - R_n(0)$ *converge uniformly to a continuous function on the interval* $[-1/2, 1/2]$ *and hence* $R_n(\theta) - R_n(0)$ *converges uniformly to a continuous function on the whole real line.*

Proof By equation (16), for $s \in [-1/2, 1/2]$

$$w_n(s) = C(s) - C(0) + E\{R_{n-1}(s+Z) - R_{n-1}(Z)\}.$$

But

$$R_{n-1}(s+Z) - R_{n-1}(Z) = f_{n-1}(s+Z) - f_{n-1}(Z) + w_{n-1}(r(s+Z)) - w_{n-1}(r(Z)).$$

Hence,

$$\begin{aligned} w_n(s) &= C(s) - C(0) + E\{f_{n-1}(s+Z) - f_{n-1}(Z)\} \\ &\quad + E\{w_{n-1}(r(s+Z)) - w_{n-1}(r(Z))\} \\ &= C(s) - C(0) + \mathcal{L}_1 f_{n-1}(s) + \mathcal{L}_2 w_{n-1}(s) \end{aligned} \tag{22}$$

where $\mathcal{L}_1$ and $\mathcal{L}_2$ are linear operators defined by

$$\begin{aligned} \mathcal{L}_1 m(s) &= \int m(x)[h(x;s) - h(x;0)]\,dx, \quad \text{and} \\ \mathcal{L}_2 m(s) &= \int m(x) Q(x,s)\,dx. \end{aligned}$$

Repeated substitution for $w_{n-1}(s)$ in (22) results in

$$w_n(s) = \sum_{i=0}^{n-1} \mathcal{L}_2^i [C(s) - C(0) + \mathcal{L}_1 f_{n-1-i}(s)]$$

where the term $\mathcal{L}_2^n w_0(s)$ has been omitted because $w_0(s) = 0$. Using the limit $f(\theta)$ from Theorem 4, this may be written as

$$w_n(s) = \sum_{i=0}^{n-1} \mathcal{L}_2^i [C(s) - C(0) + \mathcal{L}_1 f(s)] + \sum_{i=0}^{n-1} \mathcal{L}_2^i [\mathcal{L}_1 (f_{n-1-i}(s) - f(s))].$$

It will be shown that the first summation converges uniformly to a continuous function on $[-1/2, 1/2]$ and the second converges uniformly to zero.

Following the argument of Crowder (1986, Proposition 4.2) for the first summation, let

$$\begin{aligned} a &= \max_{|s| \le 1/2} [C(s) - C(0) + \mathcal{L}_1 f(s)], \\ b &= \max_{|s| \le 1/2} \max\{1, I^1(s), \ldots, I^{m-1}(s)\}, \quad \text{and} \\ \rho_0 &= \max_{|s| \le 1/2} I^m(s) \in (0,1). \end{aligned}$$

Then on $[-1/2, 1/2]$,

$$\left|\mathcal{L}_2^i[C(s) - C(0) + \mathcal{L}_1 f(s)]\right| \leq ab\rho_0^{[i/m]}$$

for $i = 0, 1, \ldots$, where $[x]$ is the greatest integer not exceeding x. Thus the terms of

$$\sum_{i=0}^{\infty} \mathcal{L}_2^i[C(s) - C(0) + \mathcal{L}_1 f(s)] \tag{23}$$

are uniformly bounded by the terms of the convergent series

$$\sum_{i=0}^{\infty} ab\rho_0^{[i/m]}$$

which assures the uniform absolute convergence of (23) on $[-1/2, 1/2]$. Since each of the terms is continuous and the convergence is uniform, the limit is continuous.

Next consider the sum

$$\sum_{i=0}^{n-1} \mathcal{L}_2^i[\mathcal{L}_1(f_{n-1-i}(s) - f(s))].$$

Using the quantities F_m, $\rho \in (0,1)$ and M_0 from the proof of Theorem 4,

$$|f_n(s) - f(s)| = \left|\sum_{m=n+1}^{\infty} F_m(0, s, 0)\right| \leq \sum_{m=n+1}^{\infty} \rho^m M_0 = \rho^n M^*$$

where $M^* = \rho M_0/(1-\rho)$. Thus,

$$|\mathcal{L}_1(f_{n-1-i}(s) - f(s))| = \left|\int [f_{n-1-i}(x) - f(x)][h(x;s) - h(x;0)]\, dx\right| \leq \rho^{n-i} M^{**}$$

where

$$M^{**} = M^* \rho^{-1} \max_{|s| \leq -1/2} \int |h(x;s) - h(x;0)|\, dx.$$

This implies

$$\begin{aligned}
\mathcal{L}_2^i[\mathcal{L}_1(f_{n-1-i}(s) - f(s))] &\leq \rho^{n-i} M^{**} b\rho_0^{[i/m]} \\
&\leq M^{**} b\rho_1^{[(n-i)/m]+[i/m]} \\
&\leq \rho_1^{-1} M^{**} b\rho_1^{[n/m]}
\end{aligned}$$

where $\rho_1 = \max\{\rho, \rho_0\} \in (0,1)$. Thus,

$$\sum_{i=0}^{n-1} |\mathcal{L}_2^i[\mathcal{L}_1(f_{n-1-i}(s) - f(s))]| \leq (\rho_1^{-1} M^{**} b) n\rho_1^{[n/m]}$$

and since $n\rho_1^{[n/m]} \to 0$ as $n \to \infty$, the series converges absolutely and uniformly to zero.

It has been shown that $w_n(s)$ converges uniformly to a continuous function on $[-1/2, 1/2]$. Since

$$R_n(\theta) - R_n(0) = f_n(\theta) + w_n(r(\theta)),$$

Theorem 4 implies the functions $R_n(\theta) - R_n(0)$ converge uniformly to a continuous function on the whole real line. □

The condition of Theorem 7 was checked numerically and found to hold at $m = 1$ for $\sigma_\nu > 0.22$ and at $m = 2$ for $\sigma_\nu > 0.19$.

6 SUMMARY AND RELEVANT LITERATURE

This paper has studied an optimal discrete adjustment problem motivated by a machining operation in the production of metal hubs. Successive hub diameters are modeled using a random walk with observation error and linear drift (corresponding to tool wear.) The tool's cutting depth is adjustable by multiples of 0.0001 inch and the problem is to determine the timing and size of adjustments. An adjustment strategy is sought to minimize expected run costs proportional to the sum of squared deviations of the hub diameters from target plus fixed costs for nonzero adjustments.

Under a simplified model omitting linear drift, optimal adjustments u_{t-1} have the form

$$\begin{aligned} u_{t-1} &= 0, && \text{if } |\hat{\theta}_{t-1}| \leq k_{n-t+1} \\ &= <-\hat{\theta}_{t-1}>, && \text{otherwise.} \end{aligned}$$

where $\hat{\theta}_{t-1}$ is a Kalman filter estimate of the process mean in period $t-1$ and $<\theta>$ is the integer nearest θ. The deadband limits k_t depend on the parameters in the probability and cost models. Plots of k_{n-t+1} versus t for various levels of adjustment resolution show that k_{n-t+1} increases near the end of the run so that one is less apt to make a costly adjustment. Section 5 proved that the limits k_{n-t+1} converge as the run length n increases.

Plots of expected run costs versus the prior mean estimate $\hat{\theta}_0$ are useful for understanding the effect of adjustment discreteness and the trade off between expected off target costs and adjustment costs. Such plots show that coarse adjustment resolution can dramatically increase expected run costs. In Section 5 the shape of the optimal expected cost functions was shown to converge. Plots of expected costs verses run length show that the incremental expected cost of including an additional period in the run is increasing but appears to converge as the horizon becomes longer.

Using the squared error plus fixed adjustment cost model Box and Jenkins (1963) studied an adjustable integrated moving average process having the same covariance structure as the random walk with observation error (and no drift.) Assuming continuous adjustments and an infinite production horizon they gave an approximation to the limiting deadband constant k_∞.

Bather (1963) studied a similar problem phrased in terms of optimal timing of machine overhauls under a random walk model with observation error (and no drift.) Although he used a rather general cost criterion, he also required (a priori) the value of a nonzero adjustment to be the negative of the estimated process mean. Under the squared error plus fixed adjustment cost model, Crowder (1986, 1991) showed that negating the estimated mean *is* the optimal nonzero (continuous) adjustment. Crowder's focus was on short production runs whereas Bather (1963) and Box and Jenkins (1963) emphasized the limiting long run problem.

Jensen (1989) incorporated linear drift and adjustment error into the probability model assuming continuous adjustments. Including a deterministic drift of d per period, shifts the deadband region [formerly $(-k_t, k_t)$] in the direction of $-d$ so it is no longer symmetric about 0. The optimal nonzero adjustment in this case has the form

$$u_{t-1} = -\hat{\theta}_{t-1} - d - z^*_{n-t+1}$$

where $-\hat{\theta}_{t-1}$ corrects for the estimated misadjustment, $-d$ corrects for drift in the next period, and $-z^*_{n-t+1}$ overcompensates in anticipation of drift which will occur in future periods before the next adjustment. Additive adjustment errors have the effect of widening the deadband region so that even without an explicit adjustment cost (i.e., even if $K_a = 0$) the optimal strategy makes no adjustment when the estimated mean is near target. In this sense the implicit cost of adding variability to the system through adjustment errors is recognized.

Kramer (1989) studied a model including costs for sampling and derived optimal sampling intervals and adjustment policies. Taguchi (1986) studied a similar problem.

An important feature of optimal adjustment problems with fixed adjustment costs is that they generally result in strategies having a deadband form wherein the process is not adjusted until a need is clearly demonstrated. This is consistent with the quality improvement philosophy which enjoins one to take corrective action only when a statistical monitor signals that a process is no longer operating in "control." In this paper and those cited above, deadband limits were derived to minimize run costs under particular probability and cost models. In contrast, "3 sigma" limits on Shewhart monitoring charts are derived to give small false alarm probabilities.

Although optimal fixed cost adjustment policies may appear to be similar to traditional Shewhart charts, it should be emphasized that they have two very different purposes. An adjustment policy is intended to regulate a process optimally. Shewhart charts were developed to to detect and allow the removal of special causes of variation not included in the nominal probability model for a process. Tucker, Faltin and Vander Wiel (1991) recommend using monitoring schemes in conjunction with feedback and feedforward adjustment policies. Combining the strengths of both methodologies allows for both short term process optimization and long term process improvement.

7 ACKNOWLEDGEMENTS

I am grateful for guidance from Stephen Vardeman as my advisor. Stephen Crowder generously supplied computer code for the continuous adjustment problem. Linda Clark helped produce the figures.

References

Bather, J. A. (1963). Control charts and the minimization of costs. *Journal of the Royal Statistical Society, Series B*, 25:49–70.

Box, G. E. P. and Jenkins, G. M. (1963). Further contributions to adaptive quality control: Simultaneous estimation of dynamics: Non-zero costs. *Bulletin of the International Statistical Institute*, 34:943–974.

Crowder, S. V. (1986). *Kalman Filtering and Statistical Process Control*. PhD thesis, Iowa State University, Ames.

Crowder, S. V. (1991). An SPC model for short production runs: Minimizing expected cost. *Technometrics*. To appear.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

Jensen, K. L. (1989). *Optimal Adjustment in the Presence of Process Drift and Adjustment Error*. PhD thesis, Iowa State University, Ames.

Kramer, T. T. (1989). *Process Control from an Economic Point of View*. PhD thesis, University of Wisconsin, Madison.

Meinhold, R. J. and Singpurwalla, N. D. (1983). Understanding the Kalman filter. *The American Statistician*, 37(2):123–127.

Taguchi, G. (1986). *Introduction to Quality Engineering: Designing Quality into Products and Processes*. Asian Productivity Organization, Tokyo.

Tucker, W. T., Faltin, F. W., and Vander Wiel, S. A. (1991). Algorithmic statistical process control: An elaboration. Technical report, Bell Labs, Murray Hill, NJ.

8 APPENDIX

Following are Lemmas 2 through 4 used in Section 3.

Lemma 2 *Let Z be a random variable with density $\phi(z)$ symmetric about $z=0$ and nonincreasing in $|z|$. Let $H(z)$ be a nonnegative function symmetric about $z=0$ such that*

$$H(z) \leq H(z+1)$$

for every $z > -1/2$. Then the function

$$G(\mu) = E\{H(\mu+Z)\}$$

is nonnegative, symmetric about $\mu=0$ and such that

$$G(\mu) \leq G(\mu+1)$$

for every $\mu > -1/2$. If in addition $\lim_{|z|\to\infty} H(z) = \infty$ and $\phi(z)$ is strictly decreasing in $|z|$, then $G(\mu) < G(\mu+1)$ for every $\mu > -1/2$.

Proof $G(\mu)$ is nonnegative since it is the expectation of the nonnegative random variable $H(\mu+Z)$. $G(\mu)$ is symmetric about $\mu=0$ since

$$\begin{aligned} G(\mu) &= \int_0^\infty [H(\mu+z)+H(\mu-z)]\phi(z)\,dz \\ &= \int_0^\infty [H(-\mu-z)+H(-\mu+z)]\phi(z)\,dz \\ &= G(-\mu). \end{aligned}$$

To show that $G(\mu) \leq G(\mu+1)$ for every $\mu > -1/2$ define the distance function

$$d(z) = |z - {<}z{>}|$$

where ${<}z{>}$ is the nearest integer function. $d(z)$ is the (positive) distance to the integer nearest z. The random variable $d(\mu+Z)$ has a density (over $\delta \in [0,1/2]$)

$$\begin{aligned} f_\mu(\delta) &= \sum_{z:d(\mu+z)=\delta} \phi(z) \\ &= \sum_{x:d(x)=\delta} \phi(x-\mu). \end{aligned}$$

It is straightforward to show that $f_{\mu+1}(\delta) = f_\mu(\delta)$. Conditioning on the random variable $d(\mu+Z)$ allows one to write

$$\begin{aligned} G(\mu) &= \int_0^\infty P[H(\mu+Z) > h]\,dh \\ &= \int_0^\infty \int_0^{1/2} \{1 - P[H(\mu+Z) \leq h | d(\mu+Z) = \delta]\} f_\mu(\delta)\,d\delta\,dh. \end{aligned} \tag{24}$$

For fixed $\delta \in [0, 1/2]$ and $h \geq 0$ define

$$\begin{aligned} p(\mu) &= P[H(\mu + Z) \leq h | d(\mu + Z) = \delta] f_\mu(\delta) \\ &= \left[\sum_{x \in A_{\delta,h}} \phi(x - \mu) \Big/ \sum_{x \in A_{\delta,\infty}} \phi(x - \mu) \right] f_\mu(\delta) \\ &= \sum_{x \in A_{\delta,h}} \phi(x - \mu) \end{aligned} \tag{25}$$

where for $0 \leq h \leq \infty$, $A_{\delta,h} = \{x : d(x) = \delta, H(x) \leq h\}$.

The sets $A_{\delta,h}$ have different forms depending on the ordering of $h, H(\delta), \lim_{i\to\infty} H(i + \delta)$ and $\lim_{i\to\infty} H(i + \delta)$. However, in each case one can demonstrate that $p(\mu) \geq p(\mu + 1)$ for every $\mu > -1/2$. Thus

$$f_\mu(\delta) - p(\mu) \leq f_\mu(\delta) - p(\mu + 1) = f_{\mu+1}(\delta) - p(\mu + 1).$$

But $f_\mu(\delta) - p(\mu)$ is the integrand in (24) and hence $G(\mu) \leq G(\mu + 1)$ for every $\mu \geq -1/2$. The last assertion is proved similarly. □

Lemma 3 *Let $H(z)$ be a function symmetric about $z = 0$ and such that*

$$H(z) \leq H(z + 1)$$

for every $z > -1/2$. If r is an element of $[-1/2, 1/2]$ and

$$\mathcal{Z}(r) = \{j + r : j = 0, \pm 1, \pm 2, \ldots\}$$

then

$$\min_{z \in \mathcal{Z}(r)} H(z) = H(r).$$

Proof The lemma follows from the inequalities

$$H(r) \leq H(1 + r) \leq H(2 + r) \leq \cdots$$

and

$$H(r) =\leq H(-1 + r) \leq H(-2 + r) \leq \cdots.$$

□

Lemma 4 *Let $F(z)$ and $H(z)$ be nonnegative functions symmetric about $z = 0$ and nondecreasing on $\{z, z + 1\}$ for every $z > -1/2$. Then* $\min\{F(z); H(z)\}$ *has these same properties.*

The proof of Theorem 3 uses the following.

Lemma 5

$$\int |h(x; \mu) - h(x; \mu + \delta)| \, dx \to 0$$

uniformly in μ as $\delta \to 0$ where $h(x; \mu)$ is the $N(\mu, \sigma_\nu^2)$ density.

Proof

$$\begin{aligned}
\int |h(x;\mu) - h(x;\mu+\delta)|\, dx &= \int |h(x;0) - h(x;\delta)|\, dx \\
&= \left| \int_{-\infty}^{\delta/2} [h(x;0) - h(x;\delta)]\, dx + \int_{\delta/2}^{\infty} [h(x;\delta) - h(x;0)]\, dx \right| \\
&= |\Phi(\delta/(2\sigma_\nu)) - \Phi(-\delta/(2\sigma_\nu)) - \Phi(-\delta/(2\sigma_\nu)) + \Phi(\delta/(2\sigma_\nu))| \\
&= 2\,|\Phi(\delta/(2\sigma_\nu)) - \Phi(\delta/(2\sigma_\nu))| \\
&\to 0
\end{aligned}$$

where Φ is the standard normal cumulative distribution function. □

Proof of Theorem 3: First 1 is shown by induction. From (11) and the definitions of $g_1(\theta)$ and $v_1(r(\theta))$ it follows that

$$\begin{aligned}
R_1(\theta) &= \min\{C(r(\theta)) + K_a; C(\theta)\} \\
&= [C(0) + K_a] + [C(r(\theta)) - C(0)] + \min\{0, C(\theta) - C(r(\theta)) - K_a\} \\
&= [C(0) + K_a] + v_1(r(\theta)) + g_1^-(\theta).
\end{aligned}$$

Suppose that for some $n \geq 1$

$$R_{n-1}(\theta) = (n-1)[C(0) + K_a] + v_{n-1}(r(\theta)) + g_{n-1}^-(\theta).$$

Using this in (11) gives

$$\begin{aligned}
R_n(\theta) &= \min\{C(r(\theta)) + K_a + (n-1)[C(0) + K_a] + \int [v_{n-1}(r(x)) + g_{n-1}^-(x)] h(x; r(\theta))\, dx; \\
&\qquad C(\theta) + (n-1)[C(0) + K_a] + \int [v_{n-1}(r(x)) + g_{n-1}^-(x)] h(x;\theta)\, dx\} \\
&= n[C(0) + K_a] \\
&\quad + \min\left\{ v_n(r(\theta))\, ; C(\theta) - C(0) - K_a + \int [v_{n-1}(r(x)) + g_{n-1}^-(x)] h(x;\theta)\, dx \right\}
\end{aligned}$$

But

$$\begin{aligned}
C(\theta) &- C(0) - K_a + \int [v_{n-1}(r(x)) + g_{n-1}^-(x)] h(x;\theta)\, dx \\
&= [C(r(\theta)) - C(0)] + \int [v_{n-1}(r(x)) + g_{n-1}^-(x)] h(x; r(\theta)) \\
&\quad + [C(\theta) - C(r(\theta)) - K_a] + \int [v_{n-1}(r(x)) + g_{n-1}^-(x)][h(x;\theta) - h(x; r(\theta))]\, dx \\
&= v_n(r(\theta)) + g_n(\theta).
\end{aligned}$$

Hence,

$$\begin{aligned}
R_n(\theta) &= n[C(0) + K_a] + \min\{v_n(r(\theta)); v_n(r(\theta)) + g_n(\theta)\} \\
&= n[C(0) + K_a] + v_n(r(\theta)) + g_n^-(\theta).
\end{aligned}$$

Next 2 is shown by induction. By the properties of $C(\theta)$ following (10),

$$g_1(\theta) = C(\theta) - C(r(\theta)) - K_a$$

is continuous, symmetric about $\theta = 0$ and strictly increasing on $\{\theta, \theta+1\}$, $\forall\theta > -1/2$.

Suppose for some $n \geq 1$ that $g_{n-1}(\theta)$ is symmetric about $\theta = 0$ and strictly increasing on $\{\theta, \theta+1\}$, $\forall\theta > -1/2$. Then $g_{n-1}^-(\theta)$ is symmetric and nondecreasing on $\{\theta, \theta+1\}$. Therefore, by Lemma 2 and since $r(\theta) = r(\theta+1)$,

$$g_n(\theta) = g_1(\theta) + \int g_{n-1}^-(x)[h(x;\theta) - h(x;r(\theta))]\,dx$$

is symmetric about $\theta = 0$ and strictly increasing on $\{\theta, \theta+1\}$, $\forall\theta > -1/2$.

To show that $g_n(\theta)$ is continuous note from part 1

$$g_{n-1}^-(\theta) - g_{n-1}^-(r(\theta)) = R_{n-1}(\theta) - R_{n-1}(r(\theta)) \equiv f_{n-1}(\theta).$$

In Lemma 1 of Section 5, $f_{n-1}(\theta)$ is shown to be uniformly bounded. Furthermore, since $g_{n-1}(\theta)$ is continuous then $g_{n-1}^-(r(\theta))$ is uniformly bounded and hence $g_{n-1}^-(\theta) = f_{n-1}(\theta) + g_{n-1}^-(r(\theta))$ is uniformly bounded by, say, $M > 0$. Next note that the symmetry of $g_{n-1}(\theta)$ implies

$$g_n(\theta) = g_1(\theta) + \int g_{n-1}^-(x)[h(x;\theta) - h(x;|r(\theta)|)]\,dx.$$

Therefore,

$$\begin{aligned}
|g_n(\theta) - g_n(\theta+\epsilon)| \;&\leq\; |g_1(\theta) - g_1(\theta+\epsilon)| \\
&\quad + \left|\int g_{n-1}^-(x)[h(x;\theta) - h(x;\theta+\epsilon)]\,dx\right| \\
&\quad + \left|\int g_{n-1}^-(x)[h(x;|r(\theta)|) - h(x;|r(\theta+\epsilon)|)]\,dx\right| \\
&\leq\; |g_1(\theta) - g_1(\theta+\epsilon)| \\
&\quad + M\int |h(x;\theta) - h(x;\theta+\epsilon)|\,dx \\
&\quad + M\int |h(x;|r(\theta)|) - h(x;|r(\theta+\epsilon)|)|\,dx \\
&\to\; 0
\end{aligned}$$

as $\epsilon \to 0$ by Lemma 5 and since $|r(\theta)|$ and $g_1(\theta)$ are continuous.

Part 3 follows directly from Theorem 1 (part 4) since the sign of $g_n(\theta)$ corresponds to the minimizer in (11). □

The final lemma was mentioned in Section 3 but developmentally follows Theorem 3 just proved.

Lemma 6 *When $C_0(\theta) = \theta^2$ the sequence $R_n(0) - R_{n-1}(0) - q_\infty$ is the same for every σ_ϵ^2.*

Proof From Theorem 3 (part 1)

$$\begin{aligned}
R_{n+1}(0) - R_n(0) - q_\infty &= [C(0) + K_a] + v_{n+1}(0) - v_n(0) - q_\infty \\
&= C(0) + K_a + \int [v_n(r(x)) - v_{n-1}(r(x))]h(x;0)\,dx \\
&\quad + \int [g_n^-(x) - g_{n-1}^-(x)]h(x;0)\,dx - q_\infty.
\end{aligned}$$

But, from the definition (10) of $C(\theta)$ it follows that

$$C(\theta) = \theta^2 + \sigma_\nu^2 + q_\infty$$

where q_∞ depends on σ_ν^2 and σ_ϵ^2. Hence,

$$\begin{aligned} R_{n+1}(0) - R_n(0) - q_\infty &= \sigma_\nu^2 + K_a + \int [v_n(r(x)) - v_{n-1}(r(x))] h(x;0)\, dx \\ &\quad + \int [g_n^-(x) - g_{n-1}^-(x)] h(x;0)\, dx. \end{aligned}$$

which does not depend on σ_ϵ^2 since neither $g_n(\theta)$ nor $v_n(r(\theta))$ do. □