A Rapid PDE-Based Optimization Methodology for Temperature Control and other Mixed Stochastic and Deterministic Systems

Sydney Howell
Finance and Accounting, Manchester Business School, Manchester, sydney.howell@mbs.ac.uk

Paul Johnson
School of Mathematics, University of Manchester, Manchester, pjohnson@maths.manchester.ac.uk

Peter Duck
School of Mathematics, University of Manchester, Manchester, duck@maths.manchester.ac.uk

Abstract

The optimal timing of temperature control, like many tasks, generally leads to analytically intractable problems: the state variables change in time, and the changes may be continuous and/or in steps, and deterministic and/or stochastic. This problem can be specified concisely as a single partial differential equation (PDE), and solved numerically within minutes (for an arbitrarily chosen control policy) on a modern PC (typically a million times faster than exhaustive simulation). By varying certain terms in the PDE, one can estimate the means and variances of many of the system’s physical and/or economic behaviors, under any chosen control policy. We also present a rapid numerical procedure for optimizing the control policy, which is robust to discontinuous problem conditions (e.g. electricity price schedules with hourly jumps); these methods have wide applicability in engineering and financial economics.
1 Introduction and literature review

Many control problems in engineering and economics are analytically intractable because they contain complex constraints and discontinuities, and a mix of stochastic and deterministic dynamics - see Björk (1998, p. 210). Problems of storage are even less tractable, whether the storage is of energy, money, or more abstract quantities. The fall-back method for intractable problems is simulation, but in many cases this is unacceptably slow or inaccurate. This paper aims to present a three-stage approach by which many continuous time-control problems can be: (i) specified as partial differential equation (PDE) systems, (ii) rapidly solved numerically under an arbitrary control policy, and (iii) rapidly solved numerically to give the optimum control policy. We use financial economic tools, e.g. Dixit and Pindyck (1994), Björk (1998) and Longstaff (2001), and extend these to an engineering problem (with potential financial implications).

In order to apply financial mathematical tools to physical control problems (in ways for which we are not aware of an exact precedent), we borrow financial mathematical concepts, but reverse their interpretation outside the mathematics. Hence we do not treat the underlying stochastic variable as a price, but as a physical variable (temperature, in our chosen heating problem). Conversely (at this stage of research) we assume the electricity price not to be stochastic, but a known function of time. We have only found one previous attempt to use a PDE to value heat storage in continuous time (in solar panels) by Haslett (1979), however his method leads to a biased solution of an unrealistically simplified problem.

Financial mathematics can directly price or optimize a system’s value without explicitly calculating the physical statistical moments (means, variances etc.) which drive the optimized value. Such physical performance parameters are often of interest, and our model can easily calculate them before or after optimization. As compared to standard financial equations, we add forcing cycles or trends to the dynamics of the stochastic variables, and deterministic dynamics for physical variables (e.g. heat flow). We also borrow a term from Asian financial options to model stock-like (integrated) variables. For optimization we bypass the well-known Hamilton-Jacobi-Bellman (HJB) equation, since this offers no direct solution method. Instead we devise a deterministic dynamic system (by adding a small auxiliary equation) whose equilibrium state is a solution of the HJB equation, and we compute this equilibrium numerically. In its fastest form this method finds the optimum decision policy at $10^7$ state points in under one minute, with a standard laptop PC.

Many well-known problems, such as those leading to Riccati equations, are special cases of our most general model (see the Appendix). In this paper we aim to present a quite general method, but for illustration we choose the simplest continuous time stochastic problem, and the simplest form of that problem, for which our model improves on existing literature, namely the optimal timing of temperature control in a building space.

The optimal timing of temperature control is driven by external temperature which varies stochastically in continuous time, with (seasonal) Brownian and mean reverting components (Yoshida and Terai, 1990; Torro, Meneu and Valor, 2003); for other models see Taylor and Buizza (2004) and Campbell and Diebold (2005). The temperature of the building space drifts towards the external temperature, unless affected by a temperature-control system. Temperature control incurs costs (at time-varying rates), but the users of a building space only suffer discomfort from undesired temperatures when they are present (e.g. for many homes, early morning and evening).

This continuous problem is often approximated in discrete time, using hourly time steps, e.g. Hsu and Su (1991), Daryanian and Norford, (1994), Henze and Krarti (1999), and also in three-hourly steps - Rolfsman (2004). Methods to optimize the approximated problem include Lagrangian relaxation (Rösch and Tröltzsch, 1992), linear programs (Henning, 1998; Ilic, Black and Watz, 2002), and a mixed integer program (Rolfsman, 2004). Alternatively, atheoretical empirical methods such as regression, neural nets and genetic algorithms can be be used e.g. Lan and Chand (1990), Papalexopoulos and Hesterberg (1990), Ren and Wright (1997), Wang and Jin (2000), Alm and Mitchell (2001), Massie (2002), Wright, Loosmore and Farmani (2002), Caldas and Norford (2003), Yalcintas and Akkurt (2003), Yang and Kim (2004) and Balasubramaniam, Samath and Kumaresan (2007). In comparing some leading optimizing algorithms (direct search, genetic algorithms and gradient search) Wetter and Wright (2004) found that gradient search fails in the face of discontinuities in the cost function - our method overcomes this problem. The
finance literature, but not so far the engineering literature, has a method which can treat optimization by simulation in continuous time, the Least Squares Monte Carlo (LSMC) method - Longstaff and Schwartz (2001). We do not pursue LSMC here because for this problem all simulation methods are far slower than our PDE method, however LSMC has advantages in the face of high dimensionality and jump statistical processes.

It seems socially optimal for end-users to pay varying electricity prices through the day, usually implemented on a fixed cycle, in which price moves hourly in discrete jumps (Ilic, Black and Watz, 2002). We model both continuous and discrete daily price cycles without instability. In future work one might specify stochastic electricity prices - see e.g. Daryanian and Norford (1994), Karakatsani (2004), Knittel and Roberts (2005), Lund and Andersen (2005), Nakamura, Nakashima and Niimura (2005), Geman and Roncoroni (2006) and Huisman, Huurman and Mahieu (2007), and with or without supplementary causal information e.g. Lo and Wu (2004) and Gonzalez, San Roque and Garcia-Gonzalez (2005).

The structure of the paper is as follows: Section 2 describes the model as a set of stochastic differential equations (SDEs), which can only be simulated at prohibitive cost in time. Section 3 combines the SDEs with economic terms in a single PDE. Section 4 solves this PDE, and optimizes its control policy, by numerical methods. Section 5 gives results, and section 6 is a summary and discussion (including of managerial implications). The Appendix gives a canonical form of our PDE, including its multivariate generalization.

2 The Model

We model climate as one variable, namely external temperature (noting that national temperature is the only traded climate variable). We model its dynamics in an accepted way, using a Brownian and a mean reverting stochastic component, plus daily seasonal forcing, c.f. Yoshida and Terai (1990) and Torro, Meneu and Valor (2003); more complexity can be added.

We model the building space itself by two parameters, its thermal capacity and its insulation; for more complex models see Hokoi and Matsumoto (1988, 1993), Kintner-Myer and Emery (1995), Fraise, Virgone and Roux (1997), Ren and Wright (1997), Braun, Montgomery and Chaturvedi (2001), Braun (2003), Tsirlin et al. (2003) and Xu et al. (2004). We assume that the user’s comfortable temperature (when present) and the insulation of the space are both constant, but deterministic time functions can easily replace the constants. We ignore the large literature which models user comfort using such extra variables as humidity, air movement etc., which have secondary effects on energy use. For time-varying models of user comfort and/or space insulation, see Oesterreicher, Bauer and Scartezzini (1996), Mozer, Vidmar and Dodier (1997) and Stuart et al. (2007).

We model the temperature control system by a single variable, namely its rate of heat import or export, with costs linear in this rate; for more complex models of such systems see e.g. Schratt, Horn and Tränkler (2000), and Wang and Jin (2000). Dynamics for the temperature control system can be added, subject to limits on dimensionality.

We have compared the costs of our optimal policy to a ‘naïve’ policy, which sets the thermostat uniformly at the ideal temperature whenever the space is occupied (morning and evening) - often called a ‘night time set back’ policy. The savings fall within the range of 10% to 30% reported in the literature for a variety of optimization or sub-optimization methods, subject to such factors as location, electricity pricing structure, building design and user occupancy schedules - see e.g. Morris, Braun and Treado (1994), Braun, Montgomery and Chaturvedi (2001) Chen (2001, 2002), and Cho and Zaheer-uddin (2003).

Under UK winter heating conditions, if the space is reasonably well insulated, our model’s optimal policy seems to save around 30% of the ‘naïve’ policy’s winter week-day electricity costs, and it halves the user’s subjective ‘cost’ of discomfort (already very low). In a badly insulated space the effects are opposite: relative to the ‘naïve’ policy, it greatly reduces discomfort, but reduces the fuel bill by only 10%. However the badly insulated space’s fuel bill is large, so the optimal policy saves more cash in the badly insulated space.
Figure 1: The temperature model: daily average. A Monte-Carlo simulation of temperature over a 6 year period. Here $T_0 = 11^\circ C$, $\alpha = 0.01$, $\sigma = 40{^\circ C/\text{year}}^{1/2}$, $a_s = 8^\circ C$, $c_s = -1560h$, $a_d = 5h$, and $c_d = 9h$.

We next model our three main state variables, external temperature, internal temperature of the space (proxied by its heat content) and the time of day.

### 2.1 Outside temperature

We assume Brownian stochastic disturbances to outside temperature, plus mean-reversion to a seasonal mean temperature, plus a forced daily cycle of temperature. Then the outside temperature, $\hat{T}$, obeys the SDE:

$$d\hat{T} = \left\{ \alpha (\hat{T}_0 + S(t) - \hat{T}) + \frac{dD}{dt} \right\} dt + \sigma dW_t,$$

where $\hat{T}_0$ is the yearly average, $S(t)$ is the seasonal component, $D(t)$ is the daily forcing component, $\alpha$ is the speed of mean reversion and $W_t$ is a Wiener process. We specify both $S(t)$ and $D(t)$ as sinusoidal functions of time, where

$$S(t) = a_s \sin\left\{ b_s (t - c_s) \right\}, \quad \text{and} \quad D(t) = a_d \sin\left\{ b_d (t - c_d) \right\},$$

where $b_s = \frac{2\pi}{8736}$ and $b_d = \frac{2\pi}{24}$. These can be generalized to other periodic, or in the case of reliable forecasts, non-periodic, functions of time.

We simulated random realizations of the above SDE using realistic parameters for typical British weather (Alaton, 2002; Brody et al., 2002) - see figure 1. Later we use this random realization to simulate the effects of an optimal heating policy through a sample day.

### 2.2 Dynamics of heat in the building space

Let $Q$ be the heat energy stored in the building space, in (for example) GWh, then if $I$ is the internal temperature at time $t$ we write, using Newton’s law of cooling

$$\frac{dQ}{dt} = k\left( C(\hat{T} - I) \right) \quad \text{or} \quad \frac{dQ}{dt} = k\left( E(\hat{T}) - Q \right),$$

where $k$ is a building space-specific constant, specifying its rate of heat loss, $C$ is a space-specific constant (with units of kWh/$^\circ C$ for example; in this paper we have taken values from Sulka and Jenkins, 2008) for thermal capacity and $E(\hat{T}) = CT$ is the equilibrium quantity of heat stored in the space when $I = \hat{T}$. 


The constant $k$ can be relaxed to a known function $k(t)$ of time, for example if doors are opened more frequently after school hours.

Now assume that the building has a heating unit, and that $H(\hat{T}, Q, t)$ is its present power output, subject to a minimum of zero and a known maximum (here usually 8KW). Any piecewise continuous hyper-surface $H(\hat{T}, Q, t)$ is a heating control policy, since the same state conditions $\hat{T}, Q, t$ cause the same heating action. Then (3) becomes

$$dQ = \left\{ H(\hat{T}, Q, t) + k\left(E(\hat{T}) - Q\right) \right\} dt. \tag{4}$$

Without loss of generality we model the problem of heating, for which $\hat{T} < I^*$; cooling is the converse problem. We next address effects on the space user.

### 2.3 Discomfort functions/prices

We define two economic costs for the space user, both in money units, namely the discomfort ‘cost’ of experiencing an undesired internal temperature, and the financial cost of temperature control. For the discomfort cost, given a desired level $I^*$ for the internal temperature, we define $Q^* = C \cdot I^*$ as the desired quantity of heat to hold in the space. When $Q$ deviates from $Q^*$ we assume that the user experiences ‘pain’ or discomfort at the rate $(Q - Q^*)^2$, and we define $p$ as the rate at which the user would be willing to spend money to eliminate this rate of discomfort, in units of (for example) dollars/[GWh]$^2$. As with $k$, the constant $p$ can be made time dependent. We define $P(t)$ as the capital value of all future expected discomfort (‘pain’) as expected at time $t$ (for simplicity we omit its dependence on $Q$ and $\hat{T}$). We assume that the user is indifferent to the temperature of the space unless physically present in it. To signal the user’s presence, the dimensionless indicator function $U(t)$ takes the value 1 when the user is present in the space and 0 at other times. Hence the time rate of change of the value $P$, which is the instantaneous rate of discomfort, in units of money per unit time, is assumed to be

$$\frac{dP}{dt} = \left\{ pU(t)(Q - Q^*)^2 \right\}. \tag{5}$$

It may not be psychologically realistic to assume step jumps between $U(t) = 0$ and $U(t) = 1$, and so we present results with and without jumps, as our estimation method can handle both. Note that by parameterizing discomfort in terms of the quantity of energy $Q$ in the space, rather than of the space’s internal temperature $I$, we make the constant $p$ unique to a given space/user combination. If needing to vary both users and spaces (which our results below mostly do not attempt) it would be useful to define $p = p^* C$ so that $p^*$ and $Q$ can be varied independently. This would define the instantaneous rate of change of capitalized discomfort $P$ as

$$\frac{dP^*}{dt} = \left\{ \frac{p^*}{C^2} U(t)(Q/C - Q^*/C)^2 \right\}, \tag{6}$$

whose units remain money per unit time; the required adjustments to the optimization criterion are straightforward. All results in this paper optimize with respect to $Q$, which may be biased towards tighter control of temperature in a larger space; the other user cost is for electrical heating.

### 2.4 Daily cycle of electricity price

The instantaneous marginal fuel cost $M$ of electricity (assuming a fixed price for fuel) depends mainly on the aggregate physical output of electricity. Progressively higher aggregate outputs use progressively older, smaller and less fuel-efficient generators, so that the ‘power stack function’ (supply curve, relating fuel usage to aggregate supply) slopes steeply upward in summer or winter evening peaks. Figure 2(a) gives an arbitrary illustration, realistic for UK conditions. In reality the supply curve is slightly stepped, to reflect the use of discrete types of generator, and the availability of each generator is stochastic. Our power stack is arbitrarily chosen to resemble effects seen in empirical data, where the spot price of electricity can double.
or triple during the peak hours of the day (Cartea and Figueroa, 2005). We use the smooth power stack to deduce a smooth mean daily schedule of electricity price, as in Figure 2(b). We later use a price schedule that moves in discrete jumps.

We define the fuel cost of electricity (per GW·h or per kW·h, for example) as $M(X, t)$ (upper curve in Figure 2(b)) where $X$ is the aggregate demand for electricity at time $t$ (lower curve in Figure 2(b)). By assuming a fixed daily cycle of aggregate electricity demand $X$, we assume that the re-timing of heating policy by end-users will not change the total electricity demand $X$, nor, therefore, its marginal cost. Under this partial equilibrium assumption the electricity price $M(X, t)$ is a function of time only, so we notate it as $M(t)$. Future work could readily explore general equilibrium solutions.

3 PDE derivation

In this section we will model, and in the subsequent section we will minimize, $V = V(\hat{T}, Q, t, H)$, which is the expected net present value (NPV) of the total cost for the space user (expressed in money) given the starting conditions $\hat{T}, Q, t$. This cost comprises: (i) money-priced discomfort, due to not enjoying the desired space temperature $T^*$ (or the desired heat quantity $Q^*$, see subsection 2.3) and (ii) the financial cost of operating a heating (cooling) system (see subsection 2.4). We write a general heating control policy as $H = H(\hat{T}, Q, t)$, the optimum heat control policy as $H^*$, and the resulting minimized expected NPV of $V$, achieved by $H^*$, as $V^* = V^*(\hat{T}, Q, t)$.

Let $V = V(\hat{T}, Q, t)$, then by Itô’s lemma the diffusive behavior of the system is

\[
\frac{dV}{dt} = \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial T^2} dt + \frac{\partial V}{\partial Q} dQ + \frac{\partial V}{\partial \hat{T}} d\hat{T} + P dt + HM dt,
\]

where all terms are cumulative flows of money during $dt$. Then inputting (1), (4), (6) we have

\[
\frac{dV}{dt} = \frac{\partial V}{\partial t} dt + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial T^2} dt + \frac{\partial V}{\partial \hat{T}} (\alpha(\hat{T}_0 + S(t) - \hat{T}) + D'(t)) dt + \sigma \frac{\partial V}{\partial \hat{T}} dW \\
+ \{H(\hat{T}, Q, t) + k(E(\hat{T}) - Q^*)\} \frac{\partial V}{\partial Q} dt + pU(t)(Q - Q^*)^2 dt + H(\hat{T}, Q, t)M(X, t) dt.
\]

We take the expectation of $dV$ (by setting the stochastic element of $d\hat{T}$ to zero) and assume that the resulting expected trend in $V$ offers the risk-free rate of return $r$ on $V$. In reality the stochastic $dV$ in any one space remains risky, but this risk is idiosyncratic (not correlated with the risky market). Therefore we can discount its expectation at the risk-free rate $r$, because over a large enough portfolio of unrelated spaces, the stochastic elements tend to sum to their mean of zero. General equilibrium work
might relax this assumption, but Cao and Wei (2000) state ‘the market price of risk associated with the [aggregate] temperature variable is insignificant in most cases’ - compare Hirschleifer and Shumway (2003).

An alternative derivation of the PDE is possible using hedging arguments, which we omit here for brevity. The assumptions above lead in a straightforward manner to the PDE:

\[
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial \hat{T}^2} + \frac{\partial V}{\partial \hat{T}} (\alpha(\hat{T}_0 + S - \hat{T}) + D') + \{H + k(E(\hat{T}) - Q)\} \frac{\partial V}{\partial Q} - rV + pU(t)(Q - Q^*)^2 + HM = 0.
\]

This PDE is a rather concise way to define the effects on the user of an arbitrary heating policy \(H\).

The first five terms on the left-hand-side of (9) are homogeneous, given \(H\), and although they have money units, they include all the physics of temperature and heat flows. However homogeneous terms alone cannot obviously trigger a non-zero value for \(V\); here the sixth and seventh terms of the PDE act as forcing terms, which enable a non-zero solution for the capital value \(V\) (here a cost).

The solution \(V(\hat{T}, Q, t)\) sums the expected NPVs of the economic forcing terms (the costs of discomfort and of fuel use) over the selected time horizon. These NPVs are discounted integrals of statistical moments, namely the first moment (expectation) of heating costs (charged as incurred) and the weighted second moment of temperature around the target (‘charged’ only at times when the user is present).

A very useful property of the PDE (9) is that by redefining the forcing terms we can capture an unbounded variety of other moments, across arbitrary subregions of the state space, under any control policy \(H\). As trivial examples, we can specialise the two forcing terms to report (for example) the costs of heating only, or the precision of temperature control in the evening period only, and so on. A more subtle application can estimate the expected time to first exit from a given region \(R\) of the state space.

We (i) define an indicator function which is unity over \(R\) (we treat the zero value outside \(R\) as a boundary condition on \(V\)), then (ii) treat the indicator as a forcing (economic) term, generating money income (coupon) at a constant rate 1. If the revised PDE is solved for \(V\), non-zero solutions exist only inside \(R\), and the solution at any point in \(R\) is the NPV of a financial instrument which will pay an annuity at the rate 1 until the system leaves \(R\). The expected time to first exit is the expected duration of this annuity, which is recovered from its capital value \(V\), its coupon 1 and the discount rate \(r\).

4 The Optimal Control Problem

To compare alternative control policies \(H\), we identify a unique heating policy as \(H_s\), where \(s\) can be a vector or scalar domain (initially the domain of \(H(\hat{T}, Q, t)\)). We must select a policy \(H_s = H^*\) over \(s\) such that when (9) is solved for \(H^* = H_s(\hat{T}, Q, t)\), the resulting value \(V_s(\hat{T}, Q, t)\) is minimized, from any starting point in \((\hat{T}, Q, t)\) space. Then we may write the requirement as

\[
V(\hat{T}, Q, t; H^*) = \min_{H \in [0,H_{\text{max}}]} V(\hat{T}, Q, t; H).
\]

The well known Hamilton-Jacobi-Bellman equation (HJB) merely states this requirement, but our problem (and many others) HJB does not define any ‘equation’ which we can directly solve to minimize \(V\), nor can it even test whether a given \(H\) minimizes \(V\). We therefore construct a dynamic system that deterministically produces the optimum \(H^*\).

It is tempting to reduce estimation effort by constraining the form of \(H^*\) to be ‘bang-bang’ rather than continuous. The ‘bang-bang’ feature exists because Brownian motion locally moves infinitely fast, but a physically realizable system can only respond at its maximum acceleration rate, hence the heating rate \(H\) should in general be at maximum or zero. In practice, for every combination of \(\hat{T}, t\) there is one level of \(Q\) above which heating is never on, which simplifies the optimal policy from a vector of (typically) \(10^7\) dimensions to a vector of (typically) \(250^2\) dimensions, still a formidable problem, even given that the solution is piecewise continuous.
However we did not exploit the simplified ‘bang-bang’ property of \( H^* \) at the start of our optimization study, for three reasons: firstly we wanted an optimization method which could, when needed, estimate a more general (piecewise) continuously variable \( H^* \). Secondly, if we specified a routine to estimate the optimum \( H^* \) in continuously variable form, a valid optimization should nonetheless estimate an \( H^* \) close to a (time-varying) ‘bang-bang’ form. This property is one of few direct tests for the optimality of the numerical solution. We indeed found that after specifying \( H \) to be continuously variable, many dynamic systems converged robustly (albeit at different speeds) to the same (time-varying) ‘bang-bang’ form of \( H^* \), and to the same \( V^* \). These estimates remained stable in face of step changes in the economic terms (e.g. hourly jumps in electricity price, as in many electricity markets). Thirdly, like Wetter and Wright (2004), when we first attempted to estimate an \( H^* \) directly in ‘bang-bang’ form the process became numerically unstable. Only after studying the continuous specification for \( H^* \) (as described below) did we find a robust method to estimate a ‘bang-bang’ form of \( H^* \). This rapidly gave solutions for \( H^* \) and \( V^* \), almost identical to those found by assuming a continuously variable \( H^* \), and was similarly robust to step changes in the occupancy \( U \) and price \( M \) functions.

We first describe our fastest method for estimating a continuous \( H^* \), since this solves the more general problem, and (in the form we used) is highly robust. We then briefly give the steps by which we reached this method, since they also suggested a fast, robust estimation method for a ‘bang-bang’ \( H^* \).

We first differentiate (9) with respect to \( H \) and set \( \phi = \frac{\partial}{\partial H} \) yielding the auxiliary problem,

\[
\frac{\partial \phi}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 \phi}{\partial T^2} + \frac{\partial \phi}{\partial T} \left( \alpha \left( T_0 + S - \hat{T} \right) + D' \right) + \left\{ H + k \left( E(T) - Q \right) \right\} \frac{\partial \phi}{\partial Q} - r \phi + \frac{\partial V}{\partial Q} + M(t) = 0. \tag{10}
\]

This PDE describes the effect of a change of the heating policy \( H \) on the option value \( V \) at the general point \((T,Q,t)\). We have the constraint \( H \in [0, h_{\text{max}}] \), but \( H \) remains continuously variable within this range, which is not specified within the PDE. In this second-order problem the function \( \phi \) can be relatively simple over the range \([0, h_{\text{max}}]\), having at most one stationary value. Therefore either \( \phi = 0 \) for some \( H^* \), which gives an interior local minimum (not ‘bang-bang’), or \( \phi > 0 \) at \( H = 0 \), or \( \phi < 0 \) at \( H = h_{\text{max}} \).

In order to force \( H \) to converge to the optimum value we need to find the zero of \( \phi \) at each state point. For this we initially guessed the \( H \) value to be identical to that at the same \((\hat{T},Q)\) point, at the previously calculated (later in time) \( t \) step (an optimal \( H \) has already been computed for that time step). This guess generally gives a non-zero \( \phi \), so we seek a correction \( \delta H \) towards \( \phi = 0 \). Newton-Raphson iteration indicates that

\[
\phi(H + \delta H) \approx \phi(H) + \delta H \frac{\partial \phi}{\partial H}. \tag{11}
\]

Then setting \( \phi(H + \delta H) = 0 \) and rearranging, the correction is \( \delta H = -\frac{\phi}{\phi'} \), where \( \phi' > 0 \) moves toward a minimum, and \( \phi' < 0 \) moves toward a maximum. In order to search for a minimum we set in this problem

\[
\delta H = -\frac{\phi}{|\phi'|}.
\]

Denoting this first guess at the optimum heating level \( H^* \), at any state point, as \( H_1 \), then we can express our next guess (the \((n+1)\)th guess, \( H_{n+1} = H(\hat{T},Q,t)_{n+1} \)) as:

\[
H_{n+1} = \min \left( H_{\text{max}}, \max \left( H_n - \frac{\phi}{|\phi'|}, 0 \right) \right). \tag{12}
\]

In this problem the coefficient \( \frac{1}{|\phi'|} \) was typically of the order of \( 10^4 \), which gave large adjustments \( \delta H \), and unusually fast convergence for a Newton-Raphson method, but it required accurate estimates of the small slope \( \phi' \). At a general time step \( t \), when solving backwards in time, we can set the differential operators in equation (10) for the previously calculated (later) \((t + dt)\)th time step to be zero, since these operators are zero if \( H \) has been optimized at step \( t + dt \). We ensure this optimality by computing the
optima backwards in time, and across the entire \((\hat{T}, Q)\) space simultaneously at each time step (c.f. the Bellman principle). Hence we can approximate \(\phi\) over \(dt\) as
\[
\phi = \frac{1}{r} \left( \frac{\partial V}{\partial Q} + M \right).
\]
The economic intuition for this simplified optimality condition (in the heating problem) is that at the optimum rate of heating \(H(\hat{T}, Q, t)\) during \(dt\), \(\phi = 0\) so the instantaneous positive total cost \(Md\), spent per unit of present heat added, equals the benefit of the fall in the NPV \(V\) of all present and future cost, which the extra heat \(dQ\) added to store will create.

### 4.1 Numerical Method

To estimate the effects of a general continuous \(H\) we use an implicit first-order finite-difference scheme. The PDE is parabolic in both \(Q\) and \(t\), so an explicit scheme for this three-dimensional problem would require restrictively small grid steps. We rejected a Crank-Nicolson scheme, because ‘bang-bang’ discontinuities in the heating policy \(H^*\), which this problem requires, cause \(V\) to oscillate in both \(Q\) and \(t\). Although such oscillations are damped for small \(dt\) or large \(t\), they obscure the accurate derivatives needed for optimization. At each time \(t\) the grid must be divided into three sections, depending on whether heat is being lost by the space or gained by it, or there is no net transfer of heat. The rate of heat loss, \(L(H, \hat{T}, Q, t)\) is defined by
\[
L(H, \hat{T}, Q, t) = H(\hat{T}, Q, t) + k_t(E(\hat{T}) - Q),
\]
which can be either negative or positive; these states imply respectively that the difference scheme must be either backwards parabolic or forwards parabolic in \(Q\).

Let \(u_{i,j,k}^t = V(i \cdot \Delta T, j \cdot \Delta Q, k \cdot \Delta \tau)\) where \(\tau = t_{\text{max}} - t\) is a simple transformation so that \(\tau\) moves backwards through time. Then for \(L < 0\), the space is cooling and we have a scheme of the following form
\[
\alpha_{i,j} u_{i-1,j}^{k+1} + \beta_{i,j} u_{i,j}^{k+1} + \gamma_{i,j} u_{i+1,j}^{k+1} + \delta_{i,j} u_{i,j-1}^{k+1} = Z_{i,j}^k, \tag{13}
\]
and for \(L > 0\) when the space is heating and we obtain
\[
\alpha_{i,j} u_{i-1,j}^{k+1} + \beta_{i,j} u_{i,j}^{k+1} + \gamma_{i,j} u_{i+1,j}^{k+1} + \delta_{i,j} u_{i,j+1}^{k+1} = Z_{i,j}^k. \tag{14}
\]
We note that the equations are generally downward parabolic in \(Q\) for smaller \(Q\), and upward parabolic in \(Q\) for larger \(Q\), implying that the solution propagates away from the center. Therefore, intriguingly, if we set \(Q_{\text{max}} \geq E(T_{\text{max}})\) and \(Q_{\text{min}} \leq E(T_{\text{min}})\) then we need not specify boundary conditions on \(V\), in the form of the exact values it must take at the maximum and minimum values of the internal temperature \(I\) (or of the internal heat content \(Q\)). There are in fact no physical limits to the maximum and minimum temperature of the space, that any practical temperature control system could approach.

For boundary conditions on \(\hat{T}\), we assume that the range of \(\hat{T}\) has been set so wide that at its limits small changes in \(\hat{T}\) are insignificant, that the first order mean reversion of the scheme dominates, so that we can omit the second-order term from (9) to obtain the modified equation
\[
\frac{\partial V}{\partial t} + \frac{\partial V}{\partial \hat{T}}(\alpha(\hat{T}_0 + S - \hat{T}) + D) + \{H + k(E(\hat{T}) - Q)\} \frac{\partial V}{\partial Q} + pU(t)(Q - Q^*)^2 + HM - rV = 0,
\]
which can be imposed at the boundary.

### 4.2 Convergence and accuracy of the scheme

It is useful to vary the sizes of the grid in the state space in order to assess accuracy. Table 1 shows that the scheme’s convergence is sufficiently smooth, such that with the aid of Richardson extrapolation, we can approximate \(V\) to around 4dp using around 320 grid steps in each direction \(3 \times 10^7\) state points. For comparison, if exhaustive simulation were to be used, the number of paths needed to estimate \(V\) to the same level of precision (also, for optimization purposes, to estimate its slope \(\phi'\) of around \(10^{-4}\) with useful accuracy) is large \((> 10^9)\) and this size of sample is needed at each of \(10^7\) state points of \(V\).
Table 1: Numerical Convergence

<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>p</th>
<th>$V(T^<em>, Q^</em>, 0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>21</td>
<td>24</td>
<td>0.6191</td>
</tr>
<tr>
<td>41</td>
<td>41</td>
<td>48</td>
<td>0.4422</td>
</tr>
<tr>
<td>81</td>
<td>81</td>
<td>96</td>
<td>0.3726</td>
</tr>
<tr>
<td>161</td>
<td>161</td>
<td>192</td>
<td>0.3330</td>
</tr>
<tr>
<td>321</td>
<td>321</td>
<td>384</td>
<td>0.3127</td>
</tr>
<tr>
<td>641</td>
<td>641</td>
<td>768</td>
<td>0.3026</td>
</tr>
</tbody>
</table>

A table to show a sample of the convergence of the scheme with numerical grid refinement for the value of discomfort over a 24 hour period in winter. $n$, $m$ and $p$ give the number of grid steps in $\hat{T}$, $Q$ and $t$ respectively. The value $V_{RE}$ is the value after first-order Richardson extrapolation.

## 5 Results

The hypersurface $V(\hat{T}, Q, t, H)$ defines the expected sum of two types of cost, discomfort and fuel, under a given control policy $H$, as integrated and discounted over a chosen time horizon. The optimal control policy $H^*$ can be plotted in three dimensions because, as noted earlier, the optimal heating rate is ‘bang-bang’ (either full on or off). Hence $H$ is either 0 or 1 (maximum) at any point in the three-dimensional $\hat{T}, Q, t$ space. The 0, 1 values are uniform over connected regions, and we can represent $H$ by the surface in the $(\hat{T}, Q, t)$ space which separates the 0 and 1 regions for $H$, noting that the hotter region is the 0 or ‘off’ region. The height of this surface in effect identifies an optimal thermostat setting $I_{th}(\hat{T}, t)$, which varies continuously with both the time of day and the external temperature - $\hat{T}$ and $I_{th}$ are continuously realized stochastically through each day. This sets simple design requirements for optimal thermostats.

The optimal policy $H^*$ minimizes the expected NPV $V$ of the sum of discomfort cost and fuel cost over any chosen time horizon, and we have noted that we can report these costs individually (likewise many other moments of financial or physical interest). In the next section we conduct a sensitivity analysis for total cost, and for its disaggregated components, discomfort and fuel cost.

### 5.1 Discounted expected value $V$ of discomfort cost, heating cost and total cost

Figure 3 assumes a fixed design of house and heating system, and varies two factors. Graph (a) assumes the optimal control policy $H^*$ and graph (b) a ‘naïve’ choice for $H$, which sets the thermostat at the desired temperature only while the space is occupied. Many heuristic strategies are known to be slightly superior to Figure 3(b), such as setting the heating on one hour before ‘entering’ the space, but the choice of which of these to display would be arbitrary. The other factor is varied continuously within each graph (a) and (b), namely the level of insulation (parameterized inversely as conductivity $k$) on the abscissa, with better insulation to the left in both plots. Resulting values are shown for total costs $V$ and for its two components, namely discomfort cost and heating cost. In each case, we summarize value across the entire cost hypersurface $V$ by a single point. This point integrates $V$ forward in time over one winter month, and over realistic ranges of the starting values of both the internal and external temperatures. Readers should note that the scale of the vertical axis differs slightly between Figure 3(a) and Figure 3(b), in order to improve discrimination within each graph (but reduces comparability between them). Under the optimal policy (a) the main cost is fuel, and discomfort cost is low. The naïve policy (b) has higher levels of both discomfort and heating costs than the optimal policy, at every level of insulation, but under the naïve policy, discomfort and heating cause similar levels of cost. Under both the optimal policy (a) and the naïve policy (b), both discomfort and heating costs rise as insulation becomes poorer (though at different rates).

We can compare absolute costs between the optimal policy (a) and the naïve policy (b). Firstly at good levels of insulation (to the left of both figures): the optimal heating cost (a) is 30% below the naïve heating
Here we show the NPV of discomfort over a winter month, for a typical house \((C = 0.8kW\cdot h/°C)\) with an 8kW heating unit. We show the total discomfort (thick solid), alongside the part made up from heating cost (solid) and user discomfort (dashed). On the left we follow the optimal strategy, and on the right a non-optimal strategy.

Cost (b) (£12 versus £18) and the optimal policy halves (a)’s already small level of discomfort. For spaces with poor levels of insulation (to the right of each graph) (a)’s optimal fuel cost is lower than (b)’s naïve cost by about 10% (£62 versus £70). This is a smaller percentage saving than in the well insulated house, but because the saving is made from a larger absolute bill, the optimal policy saves more money in the poorly insulated house than in the well insulated (£8 versus £6). The optimal policy slashes discomfort cost in the poorly insulated house from £75 to £4 through earlier but cheaper heating.

5.2 The optimal heating policy - a time-dependent optimal thermostat setting

Because the optimal heating policy \(H^*(\hat{T},Q,t)\) is ‘bang-bang’, it defines an optimal thermostat setting, an internal temperature \(I_{th}(\hat{T},t)\) below which the heating is always fully on. It is possible to plot \(I_{th}\) in 3D as a function of the (randomly evolving) external temperature \(\hat{T}\) and the time of day \(t\), but for the parameters assumed here, the effects of time of day mostly outweigh the random element of \(\hat{T}\). Therefore the effects of outside temperature are seen more clearly in a 2D plot (Figure 4) of \(I_{th}\) versus \(t\), for a selected range of discrete levels of \(\hat{T}\). It must however be stressed that a plotted contour of uniform external temperature \(\hat{T}\) does not represent a physically realizable event, since \(\hat{T}\) changes continually, both cyclically and stochastically.

Figure 4 shows the optimal thermostat setting \(I_{th}(\hat{T},t)\) through one daily cycle, for a range of outside temperatures. It illustrates optimal policies for two dwelling spaces: a house (a) and a small apartment (b). The two have identical heating systems and identical proportional rates of heat loss \(k\) but different volumes. The two optimal policies are strikingly similar in general form, but in the house the heating comes on earlier and stays on later. These differences are partly caused by the larger thermal capacity of the house, and its larger absolute heat loss (for the same \(k\) and heater power). Perhaps also our decision to measure discomfort costs in terms of quantity of heat, rather than temperature, might motivate finer control of temperature in the house, somewhat offset by the house’s generally higher fuel costs. The optimal thermostat settings change relatively smoothly, in time and in outside temperature, and there are signs of estimation instability, namely oscillation in regions where the exact thermostat setting has little effect. Smaller grid steps in time can smooth the solution surface, but with no appreciable change in \(V\), and of course require longer calculations. In reviewing these plots, it is important to recall that by the dynamics of the problem, the time profile of the thermostat setting is not identical to the time pattern of heating operation, or the time pattern of the internal temperature, and we will exemplify some of these differences later. These dynamic interactions are of course correctly specified in the PDE, and allowed for in its optimization. With this note of caution, we see that to avoid evening peak electricity prices, the
The optimal thermostat setting \( I_{th}(\hat{T}, t) \) for a dwelling with realistic marginal electricity prices, for an average household with heating power of 8kW; \( t = 0 \) and 24 correspond to midnight. The settings are shown over time for 5 different outside temperatures of \(-5^\circ C\) (highest plot), \(0^\circ C\), \(5^\circ C\), \(10^\circ C\) and \(15^\circ C\) (lowest). (a) Here the space is a house, with \( C = 0.8 \) and \( k = 0.1 \). (b) Here the space is apartment with \( C = 0.4kWh/^\circ C \) and \( k = 0.1 \).

thermostat for both dwellings is raised in three main pulses, rather than the two pulses that would match the two main periods of user activity; this is especially true when the external temperature is low.

In the first pulse, both dwellings begin heating, at low prices, long before the user becomes active, thus ensuring that the space is at or around the optimum temperature as the user becomes active. The heating tails off in anticipation of the user’s leaving the space, so as not to waste heat.

The second heating pulse, in both dwellings, starts long before the user returns, to take advantage of cheap early afternoon electricity: the house, Figure 4(a) may start heating at around 1.00 pm, depending on the outside and inside temperatures - this is some four hours before the user returns. As in the morning, the apartment starts heating an hour later than the house, and it heats more steeply. Both dwellings reach similar peak thermostat settings (the highest of the daily cycle) at 3.00 pm, and both reduce their thermostat settings fairly sharply at 4.30 pm (half an hour before the user returns) through 6.00 pm. This reduces heating at peak early evening electricity prices (see next section). Finally both dwellings raise their thermostat settings in a third main heating pulse, from 6.00 pm through 8.00 pm, when prices are falling.

Both dwellings, particularly the apartment, show signs of a very small early heating ‘spike’ at around 5.00 am. This can trigger a brief burst of early heating at lower outside temperatures, which merges into the main burst when inside temperatures are very low. This is because very early heating, although at lower prices, suffers large losses through the insulation, which is (only) justified when a large enough increase in \( Q \) is optimal before the user rises, relative to available heating power. At higher temperatures the optimal policy briefly interrupts heating, balancing higher prices against slower rates of heat loss, for smaller required delta in \( Q \).

5.3 Simulating the optimal policy’s effects in the time domain

The \( V^* \) solution produced by the optimal control policy contains no information on any actual realized time path of internal or external temperature or heating activity, since \( V^* \) at each \((Q, \hat{T}, t)\) point has integrated expected discounted cost over all possible continuous paths starting from that point, out to the time horizon used for solution. However users (and heating systems) only ever experience unique realized paths in time, therefore it is useful to see examples of the time paths which the optimal policy generates, and of the interactions which they create between \( t, \hat{T}, I_{th}, H, I \) and \( M \). For this we simulated a day’s realized time path for the external temperature (upper graph on Figure 5). The lower graph shows the realized time path of internal temperature \( I \), if the optimal policy \( H^* \) is being followed, along with the constant desired reference temperature of 21°C, which is a target only whilst the user is in the space, and
Figure 5: Simulated internal and external temperatures for the space.

A Monte-Carlo simulation for internal and external temperatures over a typical winter week for a typical house ($C = 0.8\text{kWh}/\text{°C}$, taken from Sulka and Jenkins, 2008) with average insulation ($k = 0.1$), and an $8\text{kW}$ heating unit ($t = 0$ and $24$ correspond to midnight. The external temperature (solid) is shown on top, with the internal temperature (solid) and the desired temperature level (dashed) shown below.
is shown dashed. The internal temperature starts at midnight (time 0) just below 15°C, and falls and rises strongly throughout the day, but it stays close to the desired 21°C during the periods of morning and evening occupation; these periods respectively are 7.00 am to 8.00 am, and 5.00 pm to 9.00 pm.

Figure 6 gives additional information; the upper half of this figure shows three plots: the constant dashed horizontal line is the user’s preferred internal temperature $I^\ast$ whilst in the space; the complex-looking dotted curve is this day’s randomly realized sequence of optimal thermostat settings $I_{th}$. This wanders through time, over a set of possibilities resembling those already seen in Figure 4. The remaining, continuous line on the upper half of Figure 6 is the internal temperature $I$ as before.

The lower part of Figure 6 presents two further curves. The continuous (dashed) curve shows the electricity fuel cost $M$ at each time of the day, and the discontinuous dark areas beneath this curve are the periods during this simulation when the heating was actually on. Hence the total area darkened in this graph represents the day’s total cost of heating. Starting at midnight (time $t = 0$) on the upper half of Figure 6), the internal temperature (solid line) falls to around 9 degrees at 5.00 am. There it meets (this day’s random realization of) the dotted line representing the optimal $T, I, t$ combination, at which to begin heating. The resulting heating pulse lasts until 6.30 am, after which the heating system keeps the temperature close to the desired level until the user departs at 9.00 am. Inspection of the dark area in the lower half of this figure shows that the heating system is relatively inactive during the more expensive period for electricity prices, when the user is in fact active. It uses only short bursts of heating to keep $I$ close to 21°C. The realized optimal thermostat setting falls a little just before the user leaves, trading discomfort against heat wastage.

Heating remains off until the realized fall of internal temperature meets the realized optimum heating profile for the afternoon, here at approximately 17°C and just before 3.00 pm. As in the morning, this starts a continuous burst of heating, at relatively low prices, ending shortly before 4.00 pm. Only at this stage of the sampled day does the optimal policy take the temperature significantly above the desired $I^\ast = 21$°C. After 4.00 pm the optimal thermostat setting dips sharply, to 18°C, which causes the internal temperature to fall to just above the comfortable level by 5.00 pm, when the user in fact returns. The realized $I$ then remains close to 21°C, and the heating remains off, until about 7.15 pm, after the evening price peak. Short, further bursts of heating (as seen in the lower part of the figure) keep the temperature comfortable through 9.00 pm when this user, slightly monastically, retires to bed. No heating is used after that, and the simulation ends at midnight with the realized $I$ a little higher than it was 24 hours earlier.

This sequence of simulated events shows many features that we expect of an optimal policy: it builds up heat at low prices, in advance of need, both in the morning and the afternoon, so it needs only trickle or zero inputs of heat during the morning and evening price peaks. Internal temperature stays close to $I^\ast$ during the two periods of user activity, but in the morning, when control is cheap, temperature stays rather close to $I^\ast$, while in the early evening, when control is costly, temperature varies more widely around $I^\ast$. Hence in minimising total subjective cost, the optimal policy $H^\ast$ also optimizes the precision of control, varying it visibly between the morning and evening periods, and also within each. For evening use, the optimal policy actually overheats the space, cheaply, in the afternoon (when the absent user cannot be made uncomfortable) then uses no heating during the costliest period. This allows $I$ to fall from a degree or two above $I^\ast$ to a degree or two below it. This subtle-seeming policy was not specified by us, but emerged from the unsupervised operation of the optimization algorithm.

To show that our estimation scheme can calculate robustly in face of both a fully discontinuous price function $M$, and a fully discontinuous user occupation function $U$, Figure 7 includes step discontinuities in both of these. To improve smoothness and stability we increase the number of time steps to 600 per day. In this example the electricity price makes upward step jumps before both the morning and the evening occupation periods, at which the optimal thermostat setting $I_{th}$ jumps sharply downward (conversely $I_{th}$ jumps sharply upward as the price jumps downward). The optimal policy anticipates upward jumps by sustained heating inputs, before both the morning and evening price jumps, in each case storing a surplus of cheaper heat (to above the most comfortable level) before the user is active to be inconvenienced. The optimal policy also makes step changes to the slope of the thermostat setting (but not its level) whenever the user enters or departs. At other times $I_{th}$ changes smoothly through each day, in a uniquely realized
Figure 6: Realized sequence of optimal thermostat settings, heating usage and temperature. Using a Monte-Carlo simulation for external temperatures we show the internal temperature and heating usage on a typical Monday in winter, for a typical house ($C = 0.8\text{kHz/}^\circ\text{C}$) with average insulation ($k = 0.1$), and an 8kW heating unit ($t = 0$ and $24$ correspond to midnight). The internal temperature (solid), the optimal thermostat settings (dotted) and the desired temperature level (dashed) are shown on top, while the heating usage (solid) and the relative electricity spot price $M$ (dot-dash) is shown below.

### 6 Summary and discussion

We have modeled an analytically intractable optimization problem, which contains a mixture of deterministic and stochastic dynamics, together with discontinuous economic functions. We have shown that the problem can be summarized by a single PDE, and rapidly solved numerically, under any arbitrary control policy. We have also shown that the control policy can be optimized rapidly, by methods which are robust to step changes of level in the economic terms (the forcing terms of the PDE) and also to step changes of slope (but not of level) in the PDE’s homogeneous terms. Our general optimization method rapidly estimates the optimal policy in continuous, from, which is approximately ‘bang-bang’ whenever appropriate. However we can directly estimate a ‘bang-bang’ policy even more rapidly, a task previously extremely difficult for ourselves and others.

We have noted that the economic (forcing) terms of the PDE define discounted integrals of the statistical moments which drive economic value. The PDE solution estimates the expected total of these integrals along all paths from each state point, and the optimization routine finds a control policy at each state point which makes the expected value stationary at every state point. Given the optimal control policy, the PDE can be re-solved with revised forcing terms, so as to generate many statistical moments. These include such physical performance parameters as the expected time to first exit from a chosen region $R$ of the state space.

The PDE system which we have applied to the present problem is a special case of a canonical model
Figure 7: Optimal thermostat settings and heating usage under step-changes in occupancy and cost. Using a Monte-Carlo simulation for external temperatures we show the internal temperature and heating usage on a typical Monday in winter, for a typical house ($C = 0.8\, \text{kWh}/{}^\circ\text{C}$) with average insulation ($k = 0.1$), and an $8\, \text{kW}$ heating unit ($t = 0$ and 24 correspond to midnight). The upper plot shows internal temperature (solid), the optimal thermostat settings (dotted) and the desired temperature level (dashed), whilst heating usage (solid) and the scheduled electricity price $M$ (dot-dash) is shown below.
which we give in the Appendix. The latter can define many familiar and new problems.

Our method generates continuously optimal decisions under uncertainty for the dynamic problem of heating (and cooling) a space that is used intermittently. The same method can devise an optimal time path of decision rules for many ‘constrained linear’ dynamic systems which suffer a mix of random and predictable disturbances, and where one or more quantities must be ‘stored’, subject to computation limits. In temperature control we optimize a ‘leaky’ store, namely of heat, but it is easy to model non-leaky stores, such as pumped hydro or tidal power storage. By generalising a ‘store’ to include any integrated quantity, and by adding and inter-relating several such quantities, we can optimize somewhat complex dynamic systems. For example we can in principle use a PDE to model a system containing a wind farm, a gas generator, intended to offset the wind fluctuations, and a battery intended to smooth out the remaining power disturbances. Subject to computation constraints, it is straightforward to enrich the dynamics of the model’s single equation. The generator itself, for example, can in principle be given separately constrained dynamics, for its sensor, its stochastic rate of burn and its resulting stochastic rate of rotation, an approach which can be applied to many industrial processes. Applications also seem possible in insurance, banking or other business operations, where the stored quantity is cash, and the store hopefully does not leak, but tends to grow exponentially, prior to stochastic disturbance, and at different rates for surpluses and deficits. A more challenging and remote possibility is applications in economic management, involving stochastic rates of consumption, income, savings, investment and asset accumulation, linked by accounting identities and physical dynamics, and leading to stochastically evolving physical production possibilities. Many application areas offer substantial rewards for extending present computational limits.

Appendix

The Generic Problem

When the variates $V, X, Q, P, L, E_n$ for a varying index $n$ are scalars, and of these $V, X, P, D, L$ and $E_n$ are all functions of the state variables $X, Q, t$, the canonical form of the equation which we have studied is

$$
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial X^2} + \frac{\partial V}{\partial X} D(X, Q, P, t) + L(X, Q, P, t) \frac{\partial V}{\partial Q} + E_1(X, Q, t) + E_2(X, Q, t) - rV = 0. \quad (15)
$$

Here $X$ is the stochastic factor, function $D$ (named for drift) describes how $X$ changes in expectation over $dt$ (as in a risk neutral world) $Q$ is a stored (integrated) variable, $P$ is the policy to be optimized, $L$, in general a constrained, non-linear function of $X$, describes how $Q$ changes over $dt$. The economic functions $E_n$ describe instantaneous cost or income flows, which may be functions of the state variables $X, Q, t$ and of the control policy $P$.

If all variables except the scalar $V$ are generalized to conformably dimensioned vectors, we have

$$
\frac{\partial V}{\partial t} + \frac{1}{2} \nabla_X V \hat{\mathbf{C}} \nabla_X V' + \nabla_X V D(X, Q, P, t) + L(X, Q, P, t) \nabla_Q V + E_1(X, Q, P, t) + E_2(X, Q, P, t) - rV = 0. \quad (16)
$$

where $\hat{\mathbf{C}}$ is the variance-covariance matrix of the Wiener processes within $X$ and $\nabla_X$ is a column vector of partial differentials with respect to $X$. $E_1$ and $E_2$ can be inner products and/or quadratic forms, whose elements can vary deterministically across the problem space. A special case of this model is the linear quadratic regulation problem (e.g. Björk, 1998, p. 210), which leads to Riccati equations. Under the restrictions on that model $L = I$, the identity, and $E_1$ and $E_2$ are both quadratic forms uniform across the problem space.

References


