Spending to Minimize the Cost of Infectious Disease

Stephen W. Salant

November 15, 2011

Investigation of an Unresolved Dispute

1. Sanders (1971) investigated the dynamic problem of spending over time to minimize the discounted sum of the social cost of an infection and the cost of treating it.
   - In his discrete-time formulation, he concluded that half-measures were never optimal: in each period, one should either spend nothing or the maximal amount.
2. Sethi (1974) reformulated Sanders problem in continuous time and reached the opposite conclusion:
   - if the infection level starts at the turnpike level, moderate spending to maintain that level is optimal.
   - Sethi (1974) gave his article the same title as Sanders'—distinguishing his work from Sanders' by the subtitle “a complete synthesis."

Cost-Minimization Problems with SIS Dynamics and Linear Treatment Costs

1. The Sanders-Sethi dispute is not only an interesting puzzle in its own right but promises to illuminate a problem which arises elsewhere in the literature on the optimal control of infectious disease.
2. At the heart of the Sanders-Sethi controversy is a curvature issue.
3. The issue pervades the literature, arising most recently in the twin articles by Rowthorn et al. (2009) and Anderson et al. (under revision for *Journal of Health Economics*)

Sanders' Cost-Minimization Problem

$$\min_{\gamma_t \in [0,b]} \sum_{t=1}^{\infty} \delta^t (Cx_t + K\gamma_t)$$

$$\begin{aligned} \text{subject to } x_{t+1} &= x_t + f(x_t) - \gamma_t x_t \\ x_1 &= \bar{x} \\ \text{where } f(x_t) &= \beta x_t (N - x_t),\ t = 1, 2 \ldots \end{aligned}$$

1. x_t is the endogenous number infected at time $t = 2, 3 \ldots$.
2. N is the exogenous size of the population—no births and no deaths.
3. C is the exogenous social cost per period per person infected.
4. K is the exogenous cost per intensity of treatment.
5. γ_t is either (a) the policy-maker's intensity of identifying the sick or of (b) treating the whole population

Curvature

1. As in "bang-bang" optimization problems, the Lagrangean of this discrete-time formulation is linear in γ_t
2. In Sethi's continuous-time analog, the Hamiltonian is again linear in the control.
3. The curvature problem becomes more obvious in Sanders' dynamic programming formulation: although the current cost is linear in γ, the discounted minimized future cost is strictly concave in γ.
4. Hence, one is minimizing a strictly concave function and the minimizer is always 0 or b.

An Interior Solution to the Necessary Conditions

1. Suppose one starts with the “turnpike” infection level $x^s = \frac{rK}{C - K\beta}$, where $r = \delta^{-1} - 1$.
2. Sethi says it is cost-minimizing to maintain this infection level: remove from the pool of infecteds exactly as many people as become newly infected in each period: $\gamma^s = \beta(N - x^s)$.
3. It is easily verified that this solves the Kuhn-Tucker conditions necessary for a program to minimize discounted cost.
4. But is this the only solution to these conditions?

A Simple Perturbation Reduces Costs Below Sethi's "Global Minimum"

1. Instead of Sethi's program of perpetually choosing γ^s, make two changes.
2. Pick any two consecutive periods. In the first (t) change the control to $\gamma^s + h$. This changes the infection level at $t + 1$ to $x^s(1 - h)$. Next, reduce γ by enough in $t + 1$ that the infection level returns subsequently to x^s.
3. Assume that we follow Sethi's program until t and follow it again after $t + 1$.
4. Hence, for $h = 0$ we have Sethi's program. For $h \neq 0$ we have a program differing from Sethi's in only two periods.

Implications of the Perturbation

1. The sum of the discounted costs in those two periods is a function of h.
2. This function is flat at $h = 0$, as the Kuhn-Tucker conditions reflect.
3. But the function is *strictly concave* in h.
4. Hence, if one changes γ_t in either direction non-locally around γ^s and restores $x_{t+2} = x^s$ by suitably changing γ_{t+1}, one creates a program with strictly lower costs than Sethi's turnpike program.

Dynamic Programming

1. To verify that the optimal program is never interior, let

$$V_{n+1}(x) = \min_{\gamma_{n+1} \in [0,b]} (Cx + K\gamma_{n+1}) + \delta V_n(x + f(x) - \gamma_{n+1} x),$$

   for $n = 0, 1, \ldots$
2. If $V_0 = 0$, the minimized cost $V_1(x) = Cx$ since spending in the final period is costly and provides no benefit.
3. $V_2(x)$ is strictly concave although kinked.
4. $V_2(x)$ is also strictly increasing in the inherited number infected under an innocuous parameter restriction.
5. Prior minimized cost functions inherit both properties
6. It follows that in every situation $\gamma = 0$ or $\gamma = b$. Half-measures are never optimal.

Intuition

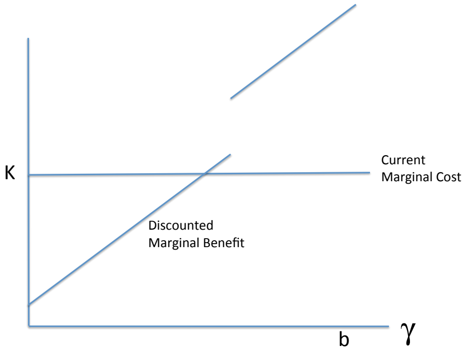

Figure: Why a Corner Solution is Always Optimal

Concluding Remarks

1. Sanders' conclusion holds in discrete time even if the length of a period is a millisecond.
2. If health authorities revise their policies periodically (every week or month or year) rather than continually, then the discrete-time formulation seems more useful.
3. However, it remains of technical interest to determine whether the advantage of Sanders' program disappears in the continuous-time limit. I'm working on that.