NOTES ON STOCHASTIC CONTROL PROBLEMS

I. ON THE STOCHASTIC LIFETIME UTILITY MAXIMIZATION PROBLEM

Nothing in life is certain: not the returns to physical or human capital investments, not the future of one's career, not even taxes. These uncertainties have a large impact on people's behavior. For instance, people work hard specifically to pay off loans, even though the traditional lifecycle models would say that only lifetime income, not the state of assets at any point in time, has any effect on behavior. People invest in several kinds of specific human capital, even though the traditional models predict perfect specialization. Such behaviors, however, can be explained as the actions of risk-averse individuals afloat in a lifetime of uncertainty.

In order to examine these kinds of behavior, one needs a lifecycle model which incorporates randomness. Unfortunately, optimal stochastic control is a messy business that cannot answer many questions. Haussmann (1986) lists eight stochastic control problems that have known closed-form solutions. Some of these have applications only in engineering. Economists who use stochastic control have focused attention on a few of these cases, especially on the LQG model: a quadratic objective, linear constraints, and Gaussian noise.

However, these certain functional forms gain closed-form solutions by omitting some effects. These specifications may miss interesting aspects of behavior because of these restrictions, and so it would be desirable to have a stochastic control technique for general functional forms.

Recent developments by Peng (1990) and Yong and Zhou (1999) allow one to derive a general maximum principle, analogous in the Itô stochastic calculus to Pontryagin's principle. There are some limitations to this technique. It is restricted to the control of Markovian processes with continuous sample paths. Perfect knowledge of the system up to and including the present is required. Technically, the principle technically applies to finite horizon problems only—existence of a solution is not clear for infinite-horizon problems. However, my arguments do not rely on whether the horizon is finite or infinite, *provided that* a solution exists and that the value function is well-defined. Finally, this solution technique allows the solution to be characterized in terms of a pair of backward-forward stochastic differential equations, which cannot be solved explicitly. However, Haussmann (1986) gives a relationship, later extended by Zhou (1996), that becomes useful. Adapting Haussman's result, I propose a solution in a manner that has a remarkably clear economic interpretation.

While these restrictions of may limit the usefulness in other applications, they serve few difficulties in lifecycle models. These principles provide an elegant and efficient way of modeling dynamics of choice under uncertainty. The solution technique allows clear economic interpretation. Moreover, it permits econometric identification of all variables of interest—we get an exact expression for the time paths of variables, and we observe how individuals respond to both the anticipated and unanticipated changes in variables.

II. THE STOCHASTIC MAXIMIZATION PRINCIPLE (CONTINUOUS TIME)

An individual seeks to maximize some objective over a predetermined interval. State variables are $x \in \mathbb{R}^n$; control variables lie in some control domain: $u \in \mathcal{U} \subseteq \mathbb{R}^m$. The objective function (often a utility function) inclusive of any discounting is $f: \mathbb{R}^n \times \mathcal{U} \times [0,\tau] \to \mathbb{R}$. The value of final states may be described by a "scrap value function" $s: \mathbb{R}^n \to \mathbb{R}$, often called a bequest motive in economics.

State variables evolve with some uncertainty. $B(\omega,t)$ is a $\mathbb{R}^k$-valued standard Wiener process on a given complete filtered probability space $\left(\Omega, \mathcal{F}, \left\{\mathcal{F}_t; t \in [0,\tau]\right\}, \mathbb{P}\right)$, where $\mathcal{F}_0$ contains all the $\mathbb{P}$–null sets in $\mathcal{F}$, and $\left\{\mathcal{F}_t; t \in [0,\tau]\right\}$ is right continuous. Evolution can be decomposed into a drift term, $g: \mathbb{R}^n \times \mathcal{U} \times [0,\tau] \to \mathbb{R}^n$, and a diffusion term, $\sigma: \mathbb{R}^n \times \mathcal{U} \times [0,\tau] \to \mathbb{R}^{n \times k}$.

Given the control space $\mathcal{U}$, an optimal control function will come from the set of adapted controls:

$$\mathcal{U}^{[0,\tau]} = \left\{ \left(u: [0,\tau] \times \Omega \to \mathcal{U}\right) : u \text{ is } \{\mathcal{F}_t\} - \text{measurable} \right\}$$

The control is said to be *admissible* when: 1.) the associated states variable $x(\cdot)$ is the unique solution to the stochastic differential equation below, in the sense that if there exists another solution $y(\cdot)$ to (9.2), then $\mathbb{P}\{\omega \in \Omega: x(\omega,t) = y(\omega,t), t \in [0,\tau]\} = 1$; and 2.) the mapping $f\left(x(\cdot),u(\cdot),\cdot\right)$ is $L^1_{\mathcal{F}}\left([0,\tau] \to \mathbb{R}\right)$.

The ***stochastic maximization problem (S-MP)*** is to maximize, over the set of admissible controls, the problem:

$$\max \mathbb{E}\left\{ \int_0^\tau f\left(x(t),u(t),t\right)dt \quad + \quad s\left(x(\tau)\right) \Big\| \{\mathcal{F}_t; t \in [0,\tau]\} \right\}$$

$$\text{subject to: } \begin{cases} dx(t) = g\left(x(t),u(t),t\right)dt + \sigma\left(x(t),u(t),t\right) \cdot dB(\omega,t) \\ x(0) = x_0 \end{cases}$$

where the stochastic differential equation (as well as the integral) is understood in the Itô sense. Solving the S-MP is analogous to solving a deterministic maximization problem, except using the Itô calculus. The ***generalized Hamiltonian*** function can be defined as:

$$\mathcal{H}\big(x(t),u(t),p(t),q(t),t\big) := f\big(x(t),u(t),t\big) + \big\langle g\big(x(t),u(t),t\big),\lambda(t)\big\rangle + \big\langle \sigma\big(x(t),u(t),t\big),\mu(t)\big\rangle$$

Whereas each constraint in the D-MP is described by a single differential equation, in the S-MP each constraint corresponds to a pair of stochastic differential equations. The $\mathcal{F}$-adapted process identified by the Kolmogorov forward equation is the first-order adjoint, $\lambda(\cdot) \in L^2_{\mathcal{F}}([0,\tau] \to \mathbb{R}^n)$. Similarly, the backward equation identifies the second-order adjoint, $\mu(\cdot) \in (L^2_{\mathcal{F}}([0,\tau] \to \mathbb{R}^n))^k$.

In full technobabble glory, the assumptions needed for the maximum principle are the following:

Assumption: *$\{\mathcal{F}_t\}$ is the natural filtration generated by B, augmented by $\mathbb{P}$–null sets in $\mathcal{F}$:*

$$\mathcal{F}_t = \sigma\big(\{B(\omega,s), s \le t\} \cup \{E \in \mathcal{F}: \mathbb{P}(E) = 0\}\big).$$

Assumption: *The control domain $\mathcal{U}$ is a convex set with nonempty interior; $\tau > 0$, so the same is true of $[0,\tau]$; and $(\mathcal{U},\rho)$ is a separable metric.*

Assumption: *The functions f, g, σ, and Γ are measurable, and there exists a constant $K > 0$ and a modulus of continuity $\overline{\omega}:[0,\infty) \to [0,\infty)$ such that the conditions: $\left|f(x,u,t) - f(x',u',t)\right| \le L\left|x - x'\right| + \overline{\omega}\big(d(u,u')\big)$, and $\left|f\big(0,u,t\big)\right| \le L$, $\left|D_x f(x,u,t) - D_x f(x',u',t)\right| \le L\left|x - x'\right| + \overline{\omega}\big(d(u,u')\big)$, and $\left|D^2_{xx} f(x,u,t) - D^2_{xx} f(x',u',t)\right| \le \overline{\omega}\big((x - x') + d(u,u')\big)$ hold $\forall t \in [0,\tau], \forall x,x' \in \mathbb{R}^n, \forall u,u' \in \mathcal{U}$ for each of the functions (replacing f with g or σ or Γ.)*

Assumption: *The functions f, g, and σ are locally Lipschitz with respect to u, and their derivatives with respect to x are continuous with respect to (x,u).*

Assumption: *The scrap value function* s *is concave, and the generalized Hamiltonian is $\mathbb{P}$-almost surely concave for all t.*

Necessary conditions for the general solution to the stochastic maximization problem was developed by Peng. They are the stochastic analogue of the Pontryagin conditions. Sufficiency was given by Zhou:

Theorem 1 (Peng's Maximum Principle, Peng 1990 and Zhou 1996): *Assume the assumptions. Let (S-MP) admit an optimal pair $\big(\overline{x}(t),\overline{u}(t)\big)$. Then the following conditions are necessary and sufficient for the quadruple $\big(\overline{x}(t),\overline{u}(t),\lambda(t),\mu(t)\big)$ to be optimal:*

1. $\big\langle D_u\mathcal{H}\big(\overline{x}(t),\overline{u}(t),\lambda(t),\mu(t),t\big),\big(u-\overline{u}\big)\big\rangle \le 0;\quad \forall u \in \mathcal{U},\ \forall t \in [0,\tau],\ \mathbb{P}-a.s.$
2. $dx(t) = D_\lambda\mathcal{H}\big(\overline{x}(t),\overline{u}(t),\lambda(t),\mu(t),t\big)dt + D_\mu\mathcal{H}\big(\overline{x}(t),\overline{u}(t),\lambda(t),\mu(t),t\big)dB(t)$
3. $d\lambda(t) = -D_x\mathcal{H}\big(\overline{x}(t),\overline{u}(t),\lambda(t),\mu(t),t\big)dt + \lambda(t)dB(t)$
4. $x(0) = x_0;\ \lambda(t) = D_x s\big(x(\tau)\big).$

Unless we know what the variables $\lambda(t)$ and $\mu(t)$ mean, these conditions are economically useless. For the sensitivity results, we need to defined the ***time-t value function***, or the maximal expected utility possible over the interval $[t,\tau]$, given that one arrives at time *t* with some states $x(t)$:

$$V(x(t),t) := \max \mathbb{E}\left\{ \int_{[t,\tau]} f(x(s),u(s),s)ds \| \{\mathcal{F}_r; r \in [t,\tau]\} \right\}$$

where the maximum is taken over all admissible controls. Concavity of the value function comes from **xxx**. If $x_i(t)$ is a "good state" (like wealth, human capital, and such) then the value function is increasing in $x_i(t)$; if it is a "bad state" (like habit formation, possibly) then the value function is decreasing in that variable. To quantify these changes, Haussman provides the usual sensitivity analysis on the value function.

Theorem 2 (Haussmann's Sensitivity Results, 1994): *Let (S1) – (S4) hold. Let (S-MP) admit an optimal pair* $(\bar{x}(t),\bar{u}(t))$. *Denote the maximized value of arriving at time* t *with states* x *by* $V(x(t),t)$, *and assume that* $V \in C^2([0,\tau]\times\mathbb{R}^n)$ *and* $D_{tx}V \in C$. *Then:*

1. $D_x V(\bar{x}(t),t) = \lambda(t); \quad \forall t \in [0,\tau], \mathbb{P}-a.s.$
2. $D_{xx}^2 V(\bar{x}(t),t)\,\sigma(\bar{x}(t),\bar{u}(t),t) = D_x\lambda(t)\,\sigma(\bar{x}(t),\bar{u}(t),t) = \mu(t); \quad \forall t \in [0,\tau], \mathbb{P}-a.s.$

If $V \in C^3([0,\tau]\times\mathbb{R}^n)$, *then both results hold* $\forall t \in [0,\tau]$.

Because this result is central to my thinking, I provide a simple derivation. It is worth noting that *a similar condition can be derived, in discrete time, using a second-order Taylor approximation rather than Itô's lemma.* Continuous time is simply much cleaner. Indirectly, this proof shows the necessity of Theorem 1.

Derivation of Theorem 2: (Borrowed from Lich-Tyler, 2001) By definition, an admissible optimal $\mathcal{F}_t$-adapted control $\bar{u}:[0,\tau]\to\mathcal{U}$ must satisfy:

$$\bar{u}:[t,\tau]\to\mathcal{U} = \arg\max \mathbb{E}\left\{ \int_{[t,\tau]} f(x(s),u(s),s)ds \| \{\mathcal{F}_r; r \in [0,\tau]\} \right\} \tag{9.15}$$

This means that u_t must almost surely satisfy:

$$\begin{aligned} V(x(t),t) &= \max_{u_t\in\mathcal{U}} \mathbb{E}_t\left\{ f(x(t),u(t),t)dt + \int_{[t+dt,\tau]} f(x(s),u(s),t)ds \| \{\mathcal{F}_r; r \in [t+dt,\tau]\} \right\} \\ &= \max_{u_t\in\mathcal{U}} \mathbb{E}_t\{ f(x(t),u(t),t)dt + V(x(t+dt),t+dt)\} \end{aligned} \tag{9.16}$$

subject to the constraint given by (9.2). I use the Itô chain rule to establish that this condition is:

$$0 = \max_{u_t} \mathbb{E}_t\left\{ f(t)dt + D_t V(t)dt + \langle dx(t), D_x V(t)\rangle + \tfrac{1}{2}\langle dx(t), D_{xx}^2 V(t)dx(t)\rangle \right\} \tag{9.17}$$

By taking the expectation, and noting that $g(x(t),u(t),t)dt$ is $\mathcal{F}_t$-predictable, while $\sigma(x(t),u(t),t)\cdot d\vec{B}(t)$ is not, I get this version of the HJB equation:

$$0 = \max_{u_t}\left\{f(t)+D_tV(t)+\left\langle g(t),D_xV(t)\right\rangle+\tfrac{1}{2}\left\langle D^2_{xx}V(t)\sigma(t),\sigma(t)\right\rangle\right\} \tag{9.18}$$

By differentiating this equation with respect to $x(t)$, we get the identity (with slight abuse of notation, for textual compactness):

$$0 = D_xf(t)+D^2_{tx}V(t)+D^2_{xx}V(t)g(t)+D_xV(t)D_xg(t)^{\mathrm{T}}+\tfrac{1}{2}D^3_{xxx}V(t)\sigma(t)^3+D^2_{xx}V(t)\sigma(t)D_x\sigma(t) \tag{9.19}$$

Rearranging, this means that:

$$\tfrac{1}{2}D^3_{xxx}V(t)\sigma(t)^2+D^2_{tx}V(t)+D^2_{xx}V(t)g(t) = -D_xf(t)-D_xV(t)D_xg(t)-D^2_{xx}V(t)\sigma(t)D_x\sigma(t) \tag{9.20}$$

What we want is an expression for the path followed by the marginal value of variable $x(t)$. Since the marginal value of $x(t)$ is $D_xV(t)$ by definition, this path is:

$$dD_xV(x(t),t) = D^2_{xx}V(x(t),t)dx(t)+\tfrac{1}{2}D^3_{xxx}V(x(t),t)\left\langle dx(t)\right\rangle+D^2_{xt}V(x(t),t)dt \tag{9.21}$$

Combining this with (9.20), we can describe the path of the marginal value of $x(t)$ as:

$$dV_x(t)-\left[D_xf(t)+\left\langle D_xg(t),D_xV(t)\right\rangle+\left\langle D_x\sigma(t),D^2_{xx}V(t)\sigma(t)\right\rangle\right]dt+D^2_{xx}V(t)\sigma(t)\cdot dB(t) \tag{9.22}$$

This path followed by $dD_xV(t)$ is the same as the path followed by $d\lambda(t)$. Since they have a common transversality condition and a common starting point, uniqueness of the solution to (9.6) allows us to infer that:

$$D_xV(\bar{x}(t),t) = \lambda(t) \tag{9.23}$$

$$D^2_{xx}V(\bar{x}(t),t)\,\sigma(\bar{x}(t),\bar{u}(t),t) = \mu(t) \tag{9.24}$$

hold almost surely for almost every $t\in[0,\tau]$. We then see that (9.18) implies (9.5), and so we have derived the maximum conditions. ■

III. The Stochastic Lifetime Utility Maximization problem (continuous time)

Enough of math for now—let's think about economics. Let us suppose that we have the (fairly general) lifetime utility maximization problem,

$$\max\mathbb{E}\left\{\int\delta(t)u(c_t)dt\,\|\{\mathcal{F}_t;t\in[0,\tau]\}\right\}$$

where c_t is time-*t* consumption, $u(\cdot)$ is the instantaneous utility function, and $\delta(t)$ is any discount function (for example, $\delta(t)=e^{-\beta t}$). Again, $\mathcal{F}_t$ can be interpreted as "information available at time *t*". I will use the shorthand $\mathbb{E}_t[x]$ to indicate $\mathbb{E}[x\|\mathcal{F}_t]$.

The evolution of assets (which can include capital, savings, money, bonds, whathaveyou) is determined by some processed. Using some functions g and σ, the mean and variance in these changes can be described.

$$\text{subject to: } \begin{cases} \mathbb{E}_t[dA_t/dt] = g(A_t, W_t, etc.) - c_t \\ \mathbb{E}_t[dA_t^2/dt] = \left(\sigma(A_t, W_t, etc.)\right)^2 \end{cases}$$

where W_t is a wage or manna-from-the-sky or something, and "*etc.*" stands for anything else that we want to stick into the function. Note that *I have not assumed any functional form for the change in assets or the utility function.* My claim is that this I express a solution without those. In fact,

Claim: *Given this stochastic lifetime utility maximization problem, the path followed by the marginal utility of wealth is:*

$$\frac{d\lambda_t/dt}{\lambda_t} = \underbrace{-\frac{\partial g_t}{\partial A_t}}_{\text{change in expectation}} - \underbrace{\frac{-\partial^2 V_t/\partial A_t^2}{\partial V_t/\partial A_t}}_{\text{absolute risk-aversion}} \cdot \underbrace{\frac{\partial \sigma_t^2}{\partial A_t}}_{\text{change in variance}} + \underbrace{\frac{-\partial^2 V_t/\partial A_t^2}{\partial V_t/\partial A_t}}_{\text{absolute risk-aversion}} \cdot \underbrace{\left(dA_t - \mathbb{E}[A_t]\right)}_{\text{unexpected income}}$$

And the consumption path is:

$$\frac{dc_t/dt}{c_t} = \eta_t \left(\underbrace{-d\ln\delta(t) - \frac{\partial g_t}{\partial A_t}}_{\text{deterministic model}} - \underbrace{\frac{-\partial^2 V_t/\partial A_t^2}{\partial V_t/\partial A_t} \cdot \frac{\partial \sigma_t^2}{\partial A_t}}_{\text{adjustment for risk-aversion}} \right) + \underbrace{\eta_t \frac{-\partial^2 V_t/\partial A_t^2}{\partial V_t/\partial A_t} \cdot \left(dA_t - \mathbb{E}[A_t]\right)}_{\text{spending unanticipated income}} + \underbrace{\pi_t \eta_t^2 \left(\frac{-\partial^2 V_t/\partial A_t^2}{\partial V_t/\partial A_t} \right)^2 \cdot \sigma_t^2}_{\text{"prudence" effect}}$$

where η_t is the intertemporal elasticity of substitution,

$$\eta_t := \frac{\partial u_t/\partial c_t}{c_t \cdot \partial u_t^2/\partial c_t^2}$$

and π_t the coefficient of relative prudence:

$$\pi_t := \frac{-c_t \cdot \partial^3 u_t/\partial c_t^3}{\partial^2 u_t/\partial h_t^2}$$

This consumption time path clearly agrees with the deterministic case, when $\sigma^2 \equiv 0$. With uncertainty, the path depends on three more ingredients: the risk in the system, the coefficient of absolute risk-aversion, and the "prudence" parameter. Note that this coefficient of risk-aversion relates to the *value function* (or lifetime utility function), rather than the *instantaneous utility function*. These rarely coincide. (I think there's a better interpretation to the lifetime one—the instantaneous one is only applicable if people absorb all of some shock within a single period. Generally, it's realistic to think that they can spread it out over the rest of their lives.)

It is easy to verify that specific cases of life-cycle consumption under uncertainty fit this formula. In cases where assets do not affect the risky component of the change in assets, when $\partial\sigma_t^2/\partial A_t \equiv 0$ (such as models with a certain real rate of return or a certain return on capital stock, see also Hall [1978]), the marginal utility of wealth indeed follows a random walk (with trend $-\partial g/\partial A_t$). When assets affect income risk, this is not necessarily true. In light of Hall's claim that the marginal utility of wealth following a random walk implies that consumption itself follows a random walk "under reasonable circumstances," we see that a necessary and sufficient circumstance for consumption to follow a random walk (even under the r.w.-MUW, a generally unlikely outcome) is that $\pi_t \equiv 0 \Leftrightarrow \partial^3 u/\partial c_t \equiv 0$, which is a quadratic utility function.[1]

I believe that the components of this solution—coefficients of absolute risk-aversion, intertemporal substitution elasticity, and prudence—have a good economic interpretation. **(Insert something about interpretation here.)** Moreover, they are econometrically identifiable.

IV. The Stochastic Lifetime Utility Maximization problem (discrete time)

Now for a discrete time version. (Not written up carefully yet. Deriving the results for the marginal utility of wealth are the same, except that I take a second-order Taylor expansion and then talk about what it looks like as $\Delta t \simeq 0$. It's the same as using Itô's lemma in the infinitesimal version. The punchline is that a problem like:

$$\max \mathbb{E}\left\{\sum \delta(t)u(c_t)\Delta t \quad \middle\| \{\mathcal{F}_t; t \in [0,\tau]\}\right\}$$

implies that:

$$\frac{\Delta\lambda_t/dt}{\lambda_t} = -\frac{\partial g_t}{\partial A_t} - \rho_t \cdot \frac{\partial\sigma_t^2}{\partial A_t} + \rho_t \cdot \left(\Delta A_t - \mathbb{E}[A_t]\right) + o(\Delta t)$$

and that:

$$\frac{\Delta c_t/dt}{c_t} = \eta_t\left(-\Delta\ln\delta(t) - \frac{\partial g_t}{\partial A_t} - \rho_t \cdot \frac{\partial\sigma_t^2}{\partial A_t}\right) + \eta_t\rho_t \cdot \left(dA_t - \mathbb{E}[A_t]\right) + \pi_t\eta_t^2\rho_t\sigma_t^2 + o(\Delta t)$$

where ρ_t stands for the coefficient of absolute risk-aversion.

---

[1] Utility functions that generate simple closed-form solutions in stochastic models, such as the quadratic utility function and the CES utility function, should be regarded with caution. A simple solution is possible for the quadratic utility function because the coefficient of prudence is identically zero. Similarly, the CES utility function restricts $\pi = \eta - 1$.

V. References

Haussmann, U.G. *A stochastic maximum principle for the optimal control of diffusions*. Pitman Research Notes in Mathematics, no. 151. Harlow, UK: Longman Scientific and Technical, 1986.

Haussmann, U.G. "Generalized solutions of the Hamilton-Jacobi equations of stochastic control." *SIAM J. of Control and Optimization* 32 (1994).

Haussmann, U.G. "Some examples of optimal stochastic control: Or, the stochastic maximum principle at work." *The SIAM Review* 23 (Jul., 1981).

Peng, S. "A general stochastic maximum principle for optimal control problems." *SIAM J. of Control and Optimization* 28 (1990).

Yong, J., and X. Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations.* New York: Springer-Verlag, 1999.

Zhou, X. Y. "A unified treatment of maximum principle and dynamic programming in stochastic controls." *Stoch. And Stoch. Rep.* 36 (1991).

Zhou, X. Y. "Sufficient conditions of optimality for stochastic systems with controllable diffusions." *IEEE Trans. Auto. Control* AC-41 (1996).