Regression Monte Carlo for solving optimal stochastic control problems

Pavel V. Shevchenko

Actuarial Studies and Business Analytics, Macquarie University, Australia

17 June 2026

Optimal Stochastic Problem

Notation:

- Controlled state variable $X^\pi = (X_t^\pi)_{t=t_0,\dots,T}$
- Control $\pi = (\pi_t)_{t=t_0,\dots,T}$
- Random disturbance $Z = (Z_t)_{t=t_0,\dots,T}$
- State variable evolution $X_{t+1}^\pi = T(X_t^\pi, \pi_t, Z_{t+1})$

Objective: maximise the expected value of the total reward

$$V_{t_0}(x) = \sup_{\pi} \mathbb{E}\left[\beta^{T-t_0} G_T(X_T^\pi) + \sum_{t=t_0}^{T-1} \beta^{t-t_0} R_t(X_t^\pi, \pi_t) | X_{t_0}^\pi = x\right],$$

where G_T and R_t are functions satisfying integrability conditions and β is discounting factor.

Optimal Stochastic Problem

This type of problem can be solved with backward recursion of the Bellman equation, where

$$\begin{aligned} V_T(x) &= G_T(x), \\ V_t(x) &= \sup_{\pi_t} \left\{ R_t(x, \pi_t) + \mathbb{E}\left[\beta V_{t+1}(X_{t+1}^{\pi}) | X_t^{\pi} = x; \pi_t \right] \right\}. \end{aligned}$$

Optimal value of control is found as

$$\pi_t^*(x) = \arg \sup_{\pi_t \in \mathcal{A}_t} \left\{ R_t(x, \pi_t) + \mathbb{E}\left[\beta V_{t+1}(X_{t+1}^{\pi}) \;\middle|\; X_t^{\pi} = x; \pi_t \right] \right\}.$$

The solution of such problem is often not possible to find analytically and numerical methods are required.
As the number of state variables increases, the numerical solution based on deterministic grid becomes very expensive computationally and simulation methods such as LSMC are favoured.

Applications of optimal stochastic control

Any application that needs to find decisions that has to be made under stochastically changing conditions, e.g. Finance, Healthcare, Engineering, Military, Climate Science and Environmental Policy, Disaster Management, etc.

My experience:

- financial planning in lifecycle and retirement: e.g. optimal purchase of annuity, optimal investment, optimal housing
- climate-economy models: e.g. optimal carbon emission reduction strategy
- Pricing and hedging financial derivatives: e.g. American/Bermudan options
- Pricing and hedging variable annuities
- Investment strategies, portfolio optimisation

Numerically, backward in time, the problem can be solved using deterministic grid in state variable space and evaluating value function between grid points via interpolation.

Algorithm Dynamic programming with deterministic grid

```
Discretize the state variable space to obtain nodes $x_j$, $j = 1, \dots, J$. This discretization can be different for different time slices, $t$. The state variable vector may include discrete and continuous variables (in this case, only continuous variables should be discretized).
Initialize $\widehat{V}_T(x_j) = 0$ for $j = 1, \dots, J$.
for $t = T - 1$ to 0 do
    Interpolate across $\widehat{V}_{t+1}(x_j)$, $j = 1, \dots, J$ to obtain the approximation $\widehat{V}_{t+1}(x)$ for any $x$. This step is not necessary when $t = T - 1$ because the maturity condition can be found for any $x$ without interpolation.
    for $j = 1, \dots, J$ do
        $\widehat{V}_t(x_j) = \sup_{\mu_t, c_t} \left( U(c_t, L_t) + e^{-\widetilde{\rho}\Delta} \mathbb{E}[\widehat{V}_{t+1}(\widetilde{x}) | \boldsymbol{X}_t = x_j] \right)$,
        where $\widetilde{x} = \mathcal{T}(x_j, \mu_t, c_t, \epsilon_{t+1})$.
    end for
end for
Simulate $M$ optimal trajectories forward in time starting from $\boldsymbol{X}_0$ to find the average, ranges, etc.
```

- In the case of one or two stochastic state variables, one can use quadrature methods to calculate expectation $\mathbb{E}[V_{t+1}(\boldsymbol{X}_{t+1})|\boldsymbol{X}_t]$. E.g. for an arbitrary function $f(x)$, the Gauss-Hermite quadrature is

$$\int_{-\infty}^{+\infty} e^{-z^2} f(z)dz \approx \sum_{j=1}^{q} W_j^{(q)} f(z_j^{(q)})$$

- In the case of many state variables and controls, simulations methods such as the regression (Least Squares) Monte Carlo simulation method are needed.

Least-Squares Monte Carlo method for high dimension models

- Least-Squares Monte Carlo (LSMC) is an approximate method for solving stochastic control problems, e.g. Longstaff and Schwartz (2001), Carriere (1996) for valuation of American options.
- Essentially a simulation and regression algorithm, where random paths are simulated and the conditional expectation in Bellman equation is approximated with a regression function, then solved via backwards recursion for stochastic control problems.
- Original exogenous LSMC extended in Kharroubi et al. (2014) with endogenous state variables and control randomisation. Adjustments are needed to LSMC in the case of utility models, see Andréasson and Shevchenko (2022).
- Notable contributions include: Tsitsiklis and Van Roy (2001), Belomestny et al. (2010); Belomestny (2011), and Aïd et al. (2014).

- If the state variable is not affected by the control, the idea behind utilitising the LSMC method is to approximate the conditional expectation

$$\Phi_t(X_t) = \mathbb{E}\left[\beta V_{t+1}(X_{t+1}) | X_t\right],$$

  by a regression scheme with independent variables X_t, and response variable $\beta V_{t+1}(X_{t+1})$. The approximation of the function is then denoted as $\widehat{\Phi}_t$.

- If the state variable is affected by control, then techniques such as control randomization are required where the conditional expectation

$$\Phi_t(X_t^\pi, \pi_t) = \mathbb{E}\left[\beta V_{t+1}(X_{t+1}^\pi) | X_t^\pi; \pi_t\right]$$

  is estimated by regression of $\beta V_{t+1}(X_{t+1})$ on X_t^π and randomised π_t Kharroubi et al. (2014). Neural network regression works better in high-dimensional case.

- For ease of notation, the superscript π on the state variable is now dropped.

Arguments for LSMC

Arguments for LSMC:

- Does not suffer from "curse of dimensionality", hence faster than other numerical methods as the number of state variables increase.
- No restrictions on dynamics of stochastic processes (contrary to PDE's). Enough to be able to simulate a path.
- Parametric estimate in feedback form of control (no grid required).

Arguments against LSMC:

- Approximate method only, and can have substantial errors piling up over multiple periods.
- Can be computationally intensive, especially for the optimisation of control variables.
- Basis function can be difficult to find and is highly problem specific - this can be resolved using deep neural networks.

Let $\mathbf{L}(X_t, \pi_t)$ be a vector of basis functions and $\mathbf{\Lambda}_t$ corresponding regression coefficients such that

$$\mathbb{E}\left[\beta V_{t+1}(X_{t+1}) | X_t; \pi_t\right] = \mathbf{\Lambda}_t' \mathbf{L}(X_t, \pi_t).$$

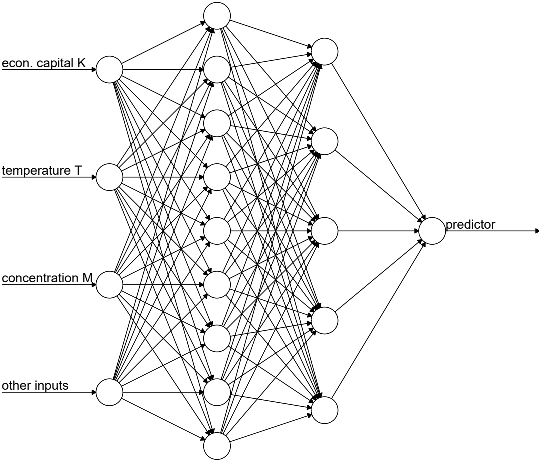
If M independent Markovian paths of state and control variables are simulated, one can consider the ordinary linear regression

$$\begin{aligned} &\beta V_{t+1}(X_{t+1}^m) = \mathbf{\Lambda}_t' \mathbf{L}(X_t^m, \pi_t^m) + \epsilon_t^m, \\ &\epsilon_t^m \overset{iid}{\sim} F_t(\cdot), \;\; \mathbb{E}[\epsilon_t^m] = 0, \;\; \mathrm{var}[\epsilon_t^m] = \sigma_t^2, \;\;\; m = 1, ..., M \end{aligned}$$

$$\hat{\mathbf{\Lambda}}_t = \arg\min_{\mathbf{\Lambda}} \sum_m \left[\beta V_{t+1}(X_{t+1}^m) - \mathbf{\Lambda}' \mathbf{L}(X_t^m, \pi_t^m)\right]^2 .$$

The above is the so-called regress now LSMC.

Typical LSMC application assumes ordinary polynomials for $\mathbf{L}(X_t, \pi_t)$ but $\boldsymbol{\Lambda}_t'\mathbf{L}(X_t, \pi_t)$ can be replaced by Neural Network approximation $\mathcal{NN}_{\boldsymbol{\theta}_t}(X_t, \pi_t)$ that requires numerical optimisation to estimate NN parameters $\boldsymbol{\theta}_t$. One can also consider transformation of response variable and modelling heteroscedasticity.



LSMC Algorithm: Endogenous state and random control

Algorithm Forward simulation

```
for $t = 0$ to $N - 1$ do
    for $m = 1$ to $M$ do
        [Simulate random samples ]
        $X_t^m := Rand \in \mathcal{X}$                                   ▷ State
        $\widetilde{\pi}_t^m := Rand \in \mathcal{A}$                      ▷ Control
        $z_{t+1}^m := Rand \in \mathcal{Z}$                                ▷ Disturbance
        [Compute the state variable after control]
        $\widetilde{X}_{t+1}^m := \mathcal{T}_t(X_t^m, \widetilde{\pi}_t^m, z_{t+1}^m)$   ▷ Evolution of state
        if $t = 0$ then $\widetilde{X}_0^m = X_0^m$
    end for
end for
```

Algorithm Backward solution (**Realised value - regress now**)

```
for $t = N$ to $0$ do
    if $t = N$ then $\widehat{V}_t(\widetilde{\boldsymbol{X}}_t) := R_N(\widetilde{\boldsymbol{X}}_t)$
    else if $t < N$ then
        [Regression of transformed value function]
        $\widehat{\boldsymbol{\Lambda}}_t := \arg\min_{\boldsymbol{\Lambda}_t} \sum_{m=1}^{M} \left[\boldsymbol{\Lambda}_t'\mathbf{L}(X_t^m, \widetilde{\pi}_t) - \beta \widehat{V}_{t+1}(\widetilde{X}_{t+1}^m)\right]^2$
        [Approximate conditional expectation] $\widehat{\Phi}_t(X_t, \widetilde{\pi}_t) := \widehat{\boldsymbol{\Lambda}}_t'\mathbf{L}(X_t, \widetilde{\pi}_t)$
        for $m = 1$ to $M$ do
            $\widehat{X}_t^m := \widetilde{X}_t^m$
            [Optimal control] $\pi_t^*(\widehat{X}_t^m) := \arg\sup_{\pi_t \in A} \left\{ R_t(\widehat{X}_t^m, \pi_t) + \widehat{\Phi}_t(\widehat{X}_t^m, \pi_t) \right\}$
            [Update value function with optimal paths]
            $\widehat{V}_t(\widehat{X}_t^m) := R_t(\widehat{X}_t^m, \pi_t^*(\widehat{X}_t^m))$
            $\widehat{X}_{t+1}^m := \mathcal{T}_t(\widehat{X}_t^m, \pi_t^*(\widehat{X}_t^m), z_t^m)$
            for $t_j = t + 1$ to $N - 1$ do
                $\widehat{V}_t(\widehat{X}_t^m) := \widehat{V}_t(\widehat{X}_t^m) + \beta^{t_j - t} R_{t_j}(\widehat{X}_{t_j}^m, \pi_{t_j}^*(\widehat{X}_{t_j}^m))$
                $\widehat{X}_{t_j+1}^m := \mathcal{T}_t(\widehat{X}_{t_j}^m, \pi_{t_j}^*(\widehat{X}_{t_j}^m), z_{t_j}^m)$
            end for
            $\widehat{V}_t(\widehat{X}_t^m) := \widehat{V}_t(\widehat{X}_t^m) + \beta^{N-t} R_N(\widehat{X}_N^m)$
        end for
    end if
end for
```

Regression surface versus realised value

There are two alternative versions of the control randomisation algorithm Kharroubi et al. (2014): the one that uses the regression surface to update the value function,

$$\widehat{V}_t(X_t) = R_t(X_t, \pi_t^*(X_t)) + \widehat{\Phi}_t(X_t, \pi_t^*(X_t)),$$

and another one that uses the realised value function,

$$\widehat{V}_t(X_t) = R_t(X_t, \pi_t^*(X_t)) + \beta \widehat{V}_{t+1}(X_{t+1}).$$

The first algorithm is the so-called value function iteration (VFI), while the second one is the so-called policy function iteration (PFI). The PFI requires recalculation of the sample paths for $t+1$ to T after each iteration backwards in time, as the optimal control affects the future state variables hence changes the simulated paths.

Overall one can do: regress now (regression surface or realised value), regress later (regression surface or realised value).

Algorithm Backward solution (**regression surface - regress later**)

```
for $t = 0$ to $N$ do
    for $m = 1$ to $M$ do
        sample $X_t^m$ in the domain of its possible values
    end for
end for
for $t = N$ to $0$ do
    if $t = N$ then
        $\widehat{V}_t(X_t) := R_N(X_t)$
    else if $t < N$ then
        [Regression/approximation of value function]
        $\widehat{\mathbf{\Lambda}}_{t+1} := \arg\min_{\mathbf{\Lambda}_{t+1}} \sum_{m=1}^{M} \left[\mathbf{\Lambda}'_{t+1}\mathbf{L}(X_{t+1}^m) - \widehat{V}_{t+1}(X_{t+1}^m)\right]^2$
        Approximate value function $\widehat{V}_{t+1}(X_{t+1}) = \widehat{\mathbf{\Lambda}}'_{t+1}\mathbf{L}(X_{t+1})$
        for $m = 1$ to $M$ do
            [Optimal control]
            $\pi_t^*(X_t^m) := \arg\sup_{\pi_t \in \mathcal{A}_t} \left\{ R_t(X_t^m, \pi_t) + \beta \mathrm{E}[\widehat{V}_{t+1}(X_{t+1}) | X_t^m; \pi_t] \right\}$
            [Update value function ]
            $\widehat{V}_t(X_t^m) := R_t(X_t^m, \pi_t^*) + \beta \mathrm{E}[\widehat{V}_{t+1}(X_{t+1}) | X_t^m; \pi_t^*]$
        end for
    end if
end for
```

Pricing Bermudan option using LSMC – numerical example

Table: Price and standard error of the Bermudan put option ($K = 40$, $r = 0.06$, $\sigma = 0.2$, $T = 1$, $N = 12$) estimated using standard LSMC ($\widehat{V}^{(0)}$), LSMC with log transformation of the value function without bias correction ($\widehat{V}^{(1)}$) and LSMC with log transformation of value function and bias correction ($\widehat{V}^{(2)}$) using *smearing estimate*. The results are based on M sample paths and 20 independent repetitions. **The 'exact' price obtained by the Binomial Tree method is $4.3862.**

M	$\widehat{V}^{(0)}$	$\widehat{V}^{(1)}$	$\widehat{V}^{(2)}$
1,000	4.4984 (0.032)	4.4336 (0.038)	4.4054 (0.039)
10,000	4.4616 (0.007)	4.4161 (0.007)	4.3962 (0.008)
100,000	4.4457 (0.003)	4.4048 (0.004)	4.3857 (0.004)

Variable annuity with GMWB

The premium paid by the policyholder upfront at t_0 is invested into the reference portfolio/risky asset $S(t)$. The value of this portfolio (**wealth account**) at time t is $W(t)$. GMWB guarantees the return of the premium via the withdrawals $\pi_n \geq 0$ allowed at times t_n, $n = 1, 2, \ldots, N$.

$$W(t_{n+1}) = \max\left(W(t_n) - \pi_n, 0\right) \frac{S(t_{n+1})}{S(t_n)} e^{-\alpha \Delta_{n+1}},$$

The guarantee balance evolves as

$$A(t_{n+1}) = A(t_n) - \pi_n \;\; n = 1, 2, \ldots, N-1, \; \pi_n \leq A(t_n)$$

The cashflow received by the policyholder at the withdrawal time t_n is given by

$$C_n(\pi_n) = \pi_n \times 1_{0 \leq \pi_n \leq G_n} + \left(G_n + (1-\beta)(\pi_n - G_n)\right) \times 1_{\pi_n > G_n},$$

where G_n is the contractual withdrawal

$$H_0(X,\pi) = \beta_{0,N} R_N(X_N) + \sum_{n=1}^{N-1} \beta_{0,n} R_n(X_n, \pi_n), \tag{1}$$

Then, the contract fair price under the given withdrawal strategy $\pi = (\pi_1, \ldots, \pi_{N-1})$, can be calculated as

$$V_0\left(X_0\right) = E\left[H_0(X,\pi)|\mathcal{F}_0\right]. \tag{2}$$

The optimal strategy is calculated as

$$\pi^*(X) = \arg\sup_{\pi\in\mathcal{A}} E\left[H_0(X,\pi)|\mathcal{F}_0\right], \tag{3}$$

$$V_n(x) = \sup_{\pi_n\in\mathcal{A}_n} \left(R_n(x,\pi_n) + E\left[\beta_{n,n+1} V_{n+1}(X_{n+1}) \middle| X_n = x; \pi_n \right] \right), \tag{4}$$

starting from the final condition $V_N(x) = R_N(x) = \max\left(W, C_N(A)\right)$.

GMWB under GBM with constant interest rate

$S(0) = 1.0$, $r = 5\%$, $g = 10\%$ ($T := 1/g = 10$ years), $\beta = 10\%$, $N_w = 1$, $\alpha = 0.0135$. All LSMC price results are averages over 20 independent LSMC runs.

Table: Prices of VA with GMWB under dynamic withdrawal strategy for different volatilities. The number of trajectories used in LSMC is $M = 1{,}000{,}000$.

$\sigma\%$	Finite Difference	OLS Algo 2, V^L	OLS Algo 3, V^U
5	0.92660 (0.00113)	0.92616 (0.00002)	0.92602 (0.00003)
10	0.94463 (0.00050)	0.94260 (0.00005)	0.94532 (0.00010)
15	0.96991 (0.00024)	0.96588 (0.00007)	0.97052 (0.00019)
20	0.99763 (0.00011)	0.99121 (0.00016)	0.99969 (0.00033)

$\sigma\%$	NN Algo 2, V^L	NN Algo 3, V^U
5	0.92406 (0.00009)	0.93320 (0.00023)
10	0.94042 (0.00016)	0.95061 (0.00014)
15	0.96386 (0.00038)	0.97749 (0.00035)
20	0.98967 (0.00034)	1.00586 (0.00048)

GMWB under GBM + Vasicek interest rate model

$$\frac{dS(t)}{S(t)} = r(t)dt + \sigma_S\left(\rho d\mathcal{B}_1(t) + \sqrt{1-\rho^2}d\mathcal{B}_2(t)\right),$$
$$dr(t) = \kappa(\theta - r(t))dt + \sigma_r d\mathcal{B}_1(t).$$

$S(0) = 1.0$, $r(0) = \theta = 5\%$, $\kappa = 0.0349$,$\sigma_r = 2\%$, $\rho = 0.3$, $g = 10\%$ ($T = 1/g$), $\beta = 10\%$, $N_w = 1$, $\alpha = 0.01$. LSMC results are the price average and corresponding standard error (in brackets next to the price estimate) over 100 independent LSMC runs. The number of trajectories is $M = 1{,}000{,}000$.

Table: Prices of VA with GMWB under dynamic withdrawal strategy in the case of Vasicek interest rate model for different volatilities.

σ%	GHQC	OLS Algo 2, V^L	OLS Algo 3, V^U	NN Algo 2, V^L	NN Algo 3, V^U
5	0.95325	0.95269 (0.005)	0.95560 (0.005)	0.95395 (0.0001)	0.95641 (0.00005)
10	0.97411	0.96976 (0.005)	0.97495 (0.005)	0.97191 (0.0002)	0.98215 (0.00009)
15	1.00147	0.99191 (0.006)	1.00500 (0.006)	0.99814 (0.0002)	1.02284 (0.00086)
20	1.03436	1.01692 (0.006)	1.03888 (0.006)	1.03748 (0.0002)	1.07160 (0.00134)

More details, see Nicolas Langrené, Xiaolin Luo, Pavel V. Shevchenko and Ruiyi Zhang (2026). Deep Least Squares Monte Carlo methods for the valuation of variable annuities with guarantees. `https://arxiv.org/abs/2605.27182`

Optimal decisions in retirement

Each period the agent receives utility, given by

$$R_t(W_t, G_t, \alpha_t, H) = \begin{cases} U_C(C_t, G_t, t) + U_H(H, G_t), & \text{if } G_t = 1, 2, \\ U_B(W_t, H), & \text{if } G_t = 0, \\ 0 & \text{if } G_t = \Delta, \end{cases}$$

with terminal condition ($t = T$) given by

$$\widetilde{R}(W_T, G_T, H) = \begin{cases} U_B(W_T, H), & \text{if } G_T \geq 0 \\ 0, & \text{if } G_T = \Delta. \end{cases}$$

We need to find a solution of the following problem

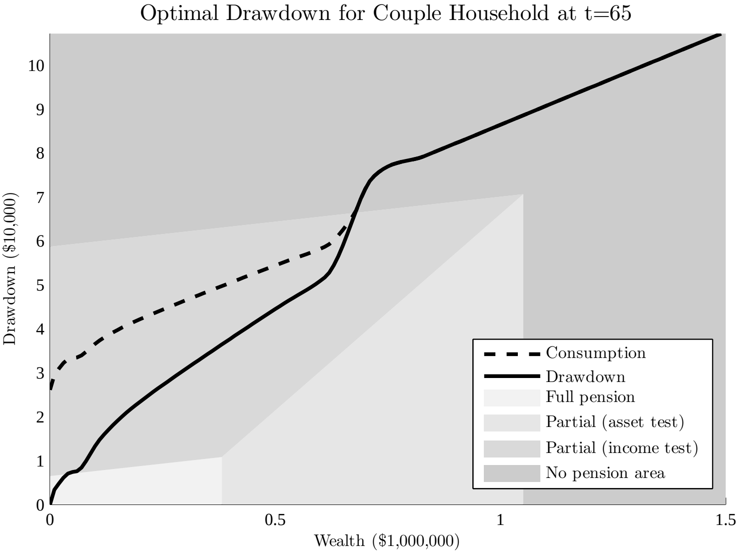
$$\widetilde{V} := \sup_{\pi} \mathbb{E}^{\pi}_{t_0} \left[\beta_{t_0,T} \widetilde{R}(W_T, G_T, H) + \sum_{t=t_0}^{T-1} \beta_{t_0,t} R_t(W_t, G_t, \alpha_t, H) \right]$$

where $\mathbb{E}^{\pi}_{t_0}[\cdot]$ is the expectation conditional on information and decision at time $t = t_0$ and $\beta_{t,t'}$ is the discounting from t to t'.

Optimal decisions in retirement

- J.G. Andréasson and P.V. Shevchenko (2022). A bias-corrected Least-Squares Monte Carlo for solving multi-period utility models. European Actuarial Journal 12, p. 349–379.
- J.G. Andréasson and P.V. Shevchenko (2024). Optimal annuitisation, housing and reverse mortgage in retirement in the presence of a means-tested public pension. European Actuarial Journal 14, p.871–904
- J. Andréasson, P. Shevchenko, and A. Novikov (2017). Optimal Consumption, Investment and Housing with Means-tested Public Pension in Retirement. Insurance: Mathematics and Economics 75, 32-47.
- J. Andréasson and P. Shevchenko (2017). Assessment of Policy Changes to Means-Tested Age Pension Using the Expected Utility Model: Implication for Decisions in Retirement. Risks 5, 47:1-47:21.

Calibration Output - Optimal Consumption



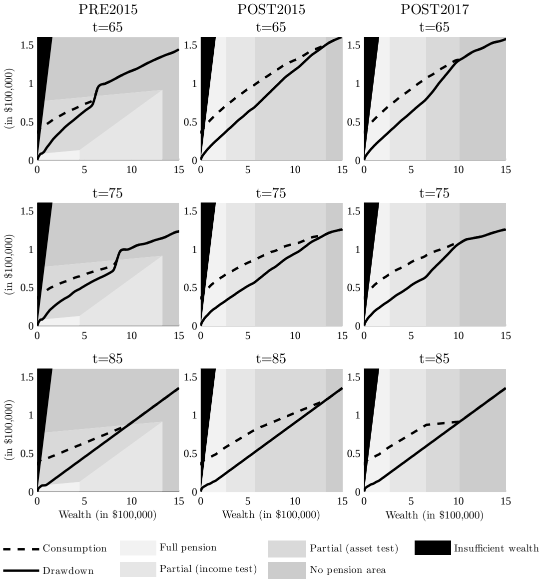


Figure: Optimal drawdown and consumption for non-homeowner couple households for a given liquid wealth at the age t, under the three different policy scenarios in the case of low returns ($\mu = 0.0325$).

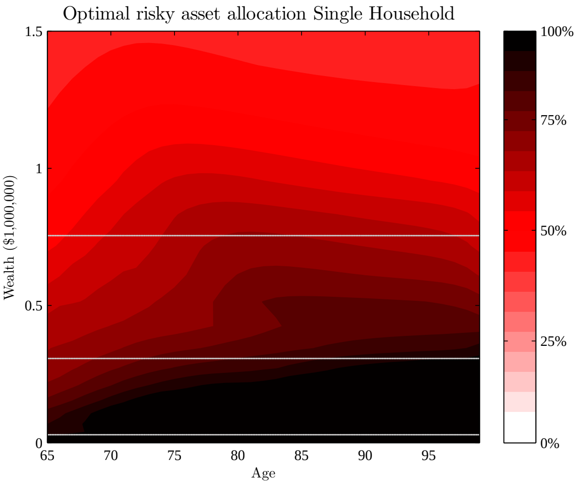


Figure: Optimal allocation of risky assets for single non-homeowners. The horizontal lines (from bottom up) show the threshold *a*, the threshold for partial Age Pension due to asset test, and the threshold for no Age Pension due to asset test.

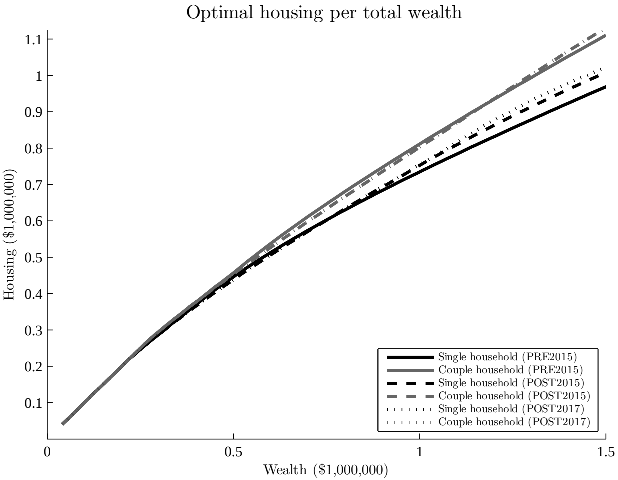


Figure: Optimal housing allocation given by total initial wealth W for single and couple households, under the three policy scenarios with the low return ($\mu = 0.0325$).

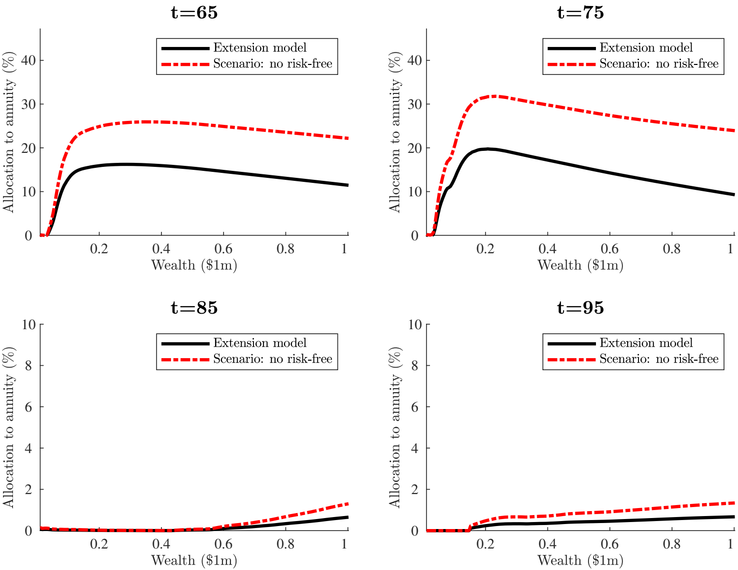


Figure: Optimal annuitisation at age t versus liquid wealth (no prior annuitisation).

The dynamic integrated climate-economy (DICE) model

$$V_0(\boldsymbol{X}_0) = \sup_{\boldsymbol{\mu},\boldsymbol{c}} \left[\sum_{t=0}^{\infty} e^{-\widetilde{\rho}\Delta t} U(c_t, L_t) \right]$$

where $U(c_t, L_t) = \frac{\Delta L_t}{1-\alpha}\left((c_t/L_t)^{1-\alpha} - 1\right)$ is utility of consumption, and state vector $\boldsymbol{X}_t = (K_t, \boldsymbol{M}_t, \boldsymbol{T}_t)$ is evolving over time as

$$\begin{aligned} K_{t+1} &= K_t(1-\delta_K)^{\Delta} + \Delta \times (Q_t(K_t, \boldsymbol{T}_t, \mu_t) - c_t) \\ \boldsymbol{M}_{t+1} &= \boldsymbol{\Phi}^M \boldsymbol{M}_t + \Delta \times (\widetilde{\beta} E_t(K_t, \mu_t), 0, 0)' \\ \boldsymbol{T}_{t+1} &= \boldsymbol{\Phi}^T \boldsymbol{T}_t + \Delta \times \left(\xi_1 F_{t+1}(M_{t+1}^{AT}), 0\right)' \end{aligned}$$

- $\boldsymbol{c} = (c_0, c_1, \ldots)$ is the consumption $c_t > 0$ and $\boldsymbol{\mu} = (\mu_0, \mu_1, \ldots)$ is the carbon emission mitigation rate $\mu_t \geq 0$.
- $K_t > 0$ is the **world economic capital**, $\boldsymbol{M}_t = (M_t^{AT}, M_t^{UP}, M_t^{LO})'$ are the **carbon concentrations in the atmosphere, upper and lower oceans**. $\boldsymbol{T}_t = (T_t^{AT}, T_t^{LO})'$ are the **temperatures in atmosphere**

$$\mathbf{\Phi}^M = \begin{pmatrix} \phi_{11} & \phi_{12} & 0 \\ \phi_{21} & \phi_{22} & \phi_{23} \\ 0 & \phi_{32} & \phi_{33} \end{pmatrix}, \ \mathbf{\Phi}^T = \begin{pmatrix} 1 - \frac{\xi_1 \eta}{\texttt{t2xco2}} - \xi_1 \xi_3 & \xi_1 \xi_3 \\ \xi_4 & 1 - \xi_4 \end{pmatrix}$$

$$\beta = 1/3.666, \ \phi_{21} = 0.12, \ \phi_{32} = 0.007$$
$$\phi_{11} = 1 - \phi_{21}, \ \phi_{33} = 1 - \phi_{23}, \ \phi_{22} = 1 - \phi_{12} - \phi_{32}$$
$$\phi_{12} = \phi_{21} \times \frac{\texttt{mateq}}{\texttt{mueq}}, \ \phi_{23} = \phi_{32} \times \frac{\texttt{mueq}}{\texttt{mleq}}$$

`mueq`=360 is equilibrium concentration in upper strata (GtC)
`mleq`=1720 is equilibrium concentration in lower strata (GtC)
`mateq`=588 is equilibrium concentration atmosphere (GtC)
`t2xco2`=3.1 is equilibrium temperature impact ($^\circ$C per doubling CO2)
$\eta = 3.6813$ is forcings of equilibrium CO2 doubling
$\xi_1 = 0.1005, \xi_3 = 0.088, \xi_4 = 0.025$ are temperature equation coefficients

- Net output: $Q_t(K_t, T_t, \mu_t) = \Omega_t(\mu_t, \sigma_t, T_t^{AT}) Y(A_t, K_t, L_t)$
- Total emission: $E_t(K_t, \mu_t) = (1-\mu_t)\sigma_t Y(A_t, K_t, L_t) + E_t^{Land}$
- Radiative forcing: $F_t(M_t^{AT}) = \eta \log_2(M_t^{AT}/\widetilde{M}^{AT}) + F_t^{EX}$

Here:

$Y(A_t, K_t, L_t) = A_t K_t^{\gamma} L_t^{1-\gamma}$: **gross GDP** (Cobb-Douglas production function)

$\Omega_t(\mu_t, \sigma_t, T_t^{AT}) = 1 - \dfrac{\sigma_t 550(1-0.025)^t \mu_t^{\theta_2}}{1000\theta_2} - \pi_2(T_t^{AT})^2$: **abatement-damage factor**

$A_t = A_{t-1}(1 + g_A(t-1))$: **total productivity factor**

$\sigma_t = \sigma_{t-1}(1 + g_\sigma(t-1))$: **decarbonisation function**

$L_t = L_{t-1}\,(11.500/L_{t-1})^{0.134}$: **world population** in billions

F_t^{EX} : **exogenous radiative forcings**

E_t^{Land} : **land emission**

- Deterministic DICE is solved using Excel Solver or GAMS by brute force solving optimisation problem wrt $\boldsymbol{c} = (c_0, \ldots, c_N)$ and $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_N)$ simultaneously

$$(\boldsymbol{\mu}^*, \boldsymbol{c}^*) = \arg\sup_{\boldsymbol{\mu}, \boldsymbol{c}} \left[\sum_{t=0}^{N} e^{-\widetilde{\rho}\Delta t} U(c_t, L_t) \right],$$

  In DICE2016: $N = 100$, $\Delta = 5$ years.

  GAMS (General Algebraic Modeling System) is a high-level programming language for mathematical modeling `https://www.gams.com/`.

- The **social cost of carbon** (USD per ton) SCC is the monetized economic loss caused by a 1-metric-ton increase in atmospheric CO_2.

$$SCC_t = -1000\widetilde{\beta} \frac{\partial V_t / \partial M_t^{\mathrm{AT}}}{\partial V_t / \partial K_t}$$

- $A(t)$, $\sigma(t)$ and L_t are exogenous functions that can be calibrated to baseline SSPs (see our paper arxiv: 2504.11721)

The stochastic DICE model

Stochastic DICE can be formulated by adding stochastic shocks to the deterministic DICE model in various places. Then we solve

$$V_0(\boldsymbol{X}_0) = \sup_{\boldsymbol{\mu},\boldsymbol{c}} \mathbb{E}\left[\sum_{t=0}^{\infty} e^{-\widetilde{\rho}\Delta t} U(c_t, L_t)\right],$$

subject to the state vector $\boldsymbol{X}_t = (K_t, \boldsymbol{M}_t, \boldsymbol{T}_t, \ldots)$ stochastically evolving over time, ie we solve optimal stochastic control problem (i.e. optimal controls at time t are functions of $\boldsymbol{X}_t$).
For example, one can consider adding shocks to state variables

$$\begin{aligned}
&K_{t+1} = K_t(1-\delta_K)^{\Delta} + \Delta \times (Q_t(K_t, T_t, \mu_t) - c_t)e^{\epsilon^K_{t+1}},\\
&\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}^M \boldsymbol{M}_t + \Delta \times (\widetilde{\beta} E_t(K_t, \mu_t), 0, 0)' e^{\epsilon^M_{t+1}},\\
&\boldsymbol{T}_{t+1} = \boldsymbol{\Phi}^T \boldsymbol{T}_t + \Delta \times \begin{pmatrix} \xi_1 F_{t+1}(M^{AT}_{t+1}) \\ 0 \end{pmatrix} + \boldsymbol{\epsilon}^T_{t+1},\ t = 0, 1, \ldots.
\end{aligned}$$

where $(\epsilon^K_t, \boldsymbol{\epsilon}^M_t, \boldsymbol{\epsilon}^T_t)$ are random disturbances.

- One can also add discrete shocks

$$I_{t+1} = \mathcal{T}^{D}(I_t, \epsilon^{I}_{t+1})$$

  e.g. affecting the gross output as

$$Y(A_t, K_t, L_t) = (1 - \chi(I_t))A_t K_t^{\gamma} L_t^{1-\gamma}$$

  where $\chi(I_t)$ is representing impact of shocks; and $I_t \in \{0, 1\}$ is representing normal and stressed regimes.
- One can add shocks to e.g. the growth rates of the output and decarbonisation leading to new state variables

$$A_{t+1} = A_t(1 + \tilde{g}^{A}_t),\ \tilde{g}^{A}_t \sim N(g^{A}_t, \sigma^2_A)$$
$$\sigma_{t+1} = \sigma_t(1 + \tilde{g}^{\sigma}_t),\ \tilde{g}^{\sigma}_t \sim N(g^{\sigma}_t, \sigma^2_\sigma)$$

- One can consider parameter uncertainty (via adding state variables) e.g. in the damage function coefficient $\hat{\pi}_2 \sim N(\pi_2, \sigma^2_\pi)$, equilibrium temperature sensitivity $\mathtt{ETS} \sim LN(t2xco2, \sigma^2_{ETS})$, etc.
- One can also consider Bayesian learning for parameters (via adding state variables for prior parameters).

SDICE can be solved using the dynamic programming

$$V_t(\boldsymbol{X}_t) = \sup_{\mu_t, c_t} \left(U(c_t, L_t) + e^{-\widetilde{\rho}\Delta} \mathbb{E}[V_{t+1}(\boldsymbol{X}_{t+1}) | \boldsymbol{X}_t] \right), \text{ s.t. } V_N(\boldsymbol{X}_N) = 0,$$

and the optimal strategy can be found as

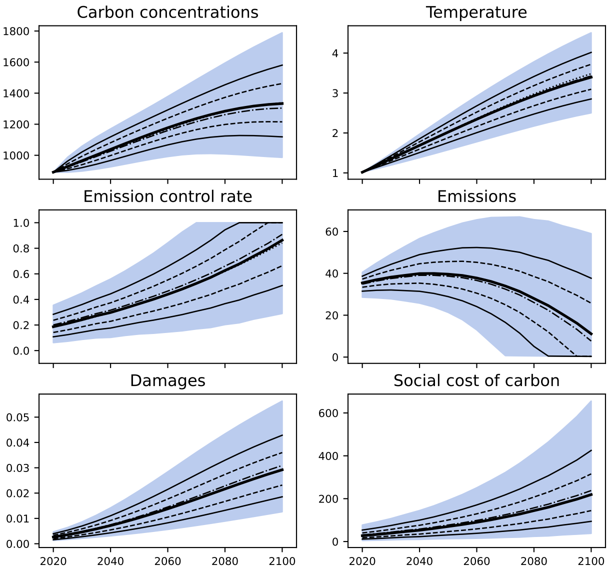
$$(\mu_t^*(\boldsymbol{X}_t), c_t^*(\boldsymbol{X}_t)) = \arg \sup_{\mu_t, c_t} \left(U(c_t, L_t) + e^{-\widetilde{\rho}\Delta} \mathbb{E}[V_{t+1}(\boldsymbol{X}_{t+1}) | \boldsymbol{X}_t] \right).$$

- A. Arandjelović, P.V. Shevchenko, T. Matsui, D. Murakami, T.A. Myrvoll (2026). Solving stochastic climate-economy models: A deep least-squares Monte Carlo approach. Mathematical Finance, 00:1-20. `https://doi.org/10.1111/mafi.70035`
- A. Arandjelović, P.V. Shevchenko (2026). The Impact of climate change on reserves in life insurance. European Actuarial Journal, https://doi.org/10.1007/s13385-026-00446-x
- P.V. Shevchenko, D. Murakami, T. Matsui, T.A. Myrvoll (2022). Impact of COVID-19 type events on the economy and climate under the stochastic DICE model. Environmental Economics and Policy Studies 24, 459-476.

Stochastic DICE via LSMC

Most significant parameter uncertainties identified in Nordhaus (2018) are equilibrium temperature sensitivity parameter, the damage function coefficient, productivity growth rate, the rate of decarbonization and carbon-cycle coefficient. Thus, for LSMC illustration we consider

- **Productivity growth rate** $A_{t+1} = A_t/(1-\tilde{g}_t^A)$, $\tilde{g}_t^A \sim N(g_t^A, \sigma_A^2)$, where $g_A(t)$ is DICE2016 growth rate, $\sigma_A = 0.056e^{-0.005t\Delta}$
- **Decarbonisation rate** $\sigma_{t+1} = \sigma_t \exp(\tilde{g}_t^\sigma \Delta)$, $\tilde{g}_t^\sigma \sim N(g_t^\sigma, \sigma_\sigma^2)$, where g_t^σ is DICE2016 growth rate, $\sigma_\sigma = 0.0032(1-0.001)^{t\Delta}$
- **Equilibrium temperature sensitivity (ETS):** $\ln ETS \sim N(1.106, 0.2646^2)$. In deterministic DICE2016, it is `t2xco2`=3.1
- **Damage function coefficient**: $\pi_2 \sim N(0.00236, (0.00236/2)^2)$
- **Carbon-cycle coefficient**: $\ln CC \sim N(5.851, 0.2649^2)$. In deterministic DICE2016, it is equilibrium concentration in upper strata `mueq`=360 (GtC)


Carbon concentrations
1800
1600
1400
1200
1000
Temperature
4
3
2
1
Emission control rate
1.0
0.8
0.6
0.4
0.2
0.0
Emissions
60
40
20
0
Damages
0.05
0.04
0.03
0.02
0.01
0.00
Social cost of carbon
600
400
200
0
2020
2040
2060
2080
2100
2020
2040
2060
2080
2100

LSMC for pricing and hedging under real process

- We can simulate underlying asset paths under the real-world probability measure rather than the conventional risk-neutral measure and find prices with hedges!
- At each discrete time step t, we simultaneously estimate the conditional continuation value $C_t(\cdot)$ of the derivative and the optimal hedge ratio $\phi_t(\cdot)$ that minimizes the hedging error/residual risk:

$$\mathcal{R}_k := \frac{1}{M}\sum_{i=1}^{M}\left(e^{-\rho}V_{t+1}^{(i)} - \widehat{C}_t(S_t^{(i)}) + \widehat{\phi}_t(S_t^{(i)})\big(S_t^{(i)} - e^{-\rho}S_{t+1}^{(i)}\big)\right)^2.$$

Continuation value is approximated by a neural network; one can calculate delta or set it via automatic differentiation:

$$\widehat{C}_t(s) \;=\; f_t(s;\theta_t),\; \widehat{\phi}_t(s) \;:=\; \frac{\partial \widehat{C}_t(s)}{\partial s}$$

This is joint work in progress with Nicolas Langrené and Wen Chen.

Table: American put pricing comparison on a non-dividend-paying stock under GBM. Columns report: (i) Longstaff (2001) finite-difference American benchmark, (ii) Black–Scholes closed-form European put, (iii) Longstaff (2001) simulated American estimate with Monte Carlo standard error (s.e.), and (iv) our NN–HMC simulated American estimate with Longstaff-style (s.e.). Here $K = 40$, $r = 0.06$, $N = 50T$, and M = 100,000 for both Longstaff and HMC exercise opportunities per year.

S_0	σ	T	N	FD American	European	Longstaff (s.e.)	NN–HMC (s.e.)
40	0.20	1	50	2.314	2.066	2.313 (0.009)	2.317 (0.009)
40	0.20	2	100	2.885	2.356	2.879 (0.010)	2.885 (0.010)
40	0.40	1	50	5.312	5.060	5.308 (0.018)	5.314 (0.018)
40	0.40	2	100	6.920	6.326	6.921 (0.022)	6.906 (0.022)

Table: Kemna–Vorst (1990) Table 1 grid replicated with our NN–HMC for an arithmetic-average (AV) Asian call. We use $S_0 = 40$, $T = \frac{1}{3}$ year, $N = 17$ monitoring dates (50 steps/year), and gross annual factor $r(\mathrm{gross})$. Standard errors (s.e.) are reported in parentheses.

$r(\mathrm{gross})$	σ	K	BS Call C^a	Std MC	(s.e.)	NN–HMC	(s.e.)
1.03	0.20	35.0	5.564171	5.159756	(0.00818)	5.160903	(0.00024)
1.03	0.20	40.0	2.035335	1.139490	(0.00526)	1.140847	(0.00085)
1.03	0.20	45.0	0.458753	0.051594	(0.00108)	0.052576	(0.00036)
1.03	0.30	35.0	6.079410	5.268310	(0.01179)	5.270904	(0.00064)
1.03	0.30	40.0	2.947493	1.657017	(0.00798)	1.658257	(0.00126)
1.03	0.30	45.0	1.185225	0.269700	(0.00326)	0.271930	(0.00088)
1.03	0.40	35.0	6.739822	5.490520	(0.01509)	5.494525	(0.00106)
1.03	0.40	40.0	3.858478	2.174043	(0.01083)	2.176325	(0.00168)
1.03	0.40	45.0	2.021782	0.614931	(0.00593)	0.618597	(0.00140)

Table: Down-and-out call (discrete monitoring $m = 50$): Broadie et al. (1997) Table 2.1 benchmark vs. NN–HMC. Parameters: $S_0 = 100$, $K = 100$, $\sigma = 0.30$, $r = 0.10$, $T = 0.2$, $N = 50$.

H	True	NN–HMC	(s.e.)	RelErr (%)
85	6.322	6.321	(0.002)	-0.01
86	6.306	6.308	(0.002)	0.03
87	6.281	6.282	(0.002)	0.01
88	6.242	6.243	(0.002)	0.01
89	6.184	6.183	(0.002)	-0.02
90	6.098	6.100	(0.002)	0.03
91	5.977	5.979	(0.002)	0.04
92	5.810	5.810	(0.002)	0.00
93	5.584	5.586	(0.002)	0.04
94	5.288	5.288	(0.002)	-0.01
95	4.907	4.909	(0.002)	0.03
96	4.427	4.428	(0.002)	0.01
97	3.834	3.832	(0.002)	-0.05
98	3.126	3.127	(0.003)	0.05
99	2.337	2.335	(0.003)	-0.11

Table: Static GMWB fair fee comparison. The paper benchmark is the Monte Carlo (MC) result reported in Luo and Shevchenko (2015), Table 1.

Contractual rate g (%)	Maturity $T = 1/g$	Paper MC (bp)	NN–HMC (bp)	Abs. Diff. (bp)
4	25.00	17.23	17.72	0.49
5	20.00	28.29	28.56	0.27
6	16.67	40.37	40.28	-0.09
7	14.29	53.20	53.17	-0.03
8	12.50	67.02	66.94	-0.08
9	11.11	81.23	81.01	-0.22
10	10.00	95.79	95.65	-0.14
15	6.67	171.50	170.65	-0.85

Concluding remarks

- Regression Monte Carlo is very powerful numerical method to solve many important and interesting problems for decisions under uncertainty across various areas.
- Combined with NN approximations, it is more powerful and easier to use.

References

Aïd, René, Luciano Campi, Nicolas Langrené, and Huyên Pham (2014), "A probabilistic numerical method for optimal multiple switching problems in high dimension." *SIAM Journal on Financial Mathematics*, 5, 191–231.

Andreasson, Johan G and Pavel V Shevchenko (2019), "Bias-corrected Least-Squares Monte Carlo for utility based optimal stochastic control problems." *Preprint SSRN: 2985828*. Available at `https://ssrn.com`.

Belomestny, Denis (2011), "Pricing Bermudan options by nonparametric regression: Optimal rates of convergence for lower estimates." *Finance and Stochastics*, 15, 655–683.

Belomestny, Denis, Anastasia Kolodko, and John Schoenmakers (2010), "Regression methods for stochastic control problems and their convergence analysis." *SIAM Journal on Control and Optimization*, 48, 3562–3588.

Carriere, Jacques (1996), "Valuation of the early-exercise price for options using simulations and nonparametric regression." *Insurance: Mathematics and Economics*, 19, 19–30.

Kharroubi, Idris, Nicolas Langrené, and H Pham (2014), "A numerical algorithm for fully nonlinear HJB equations: an approach by control randomization." *Monte Carlo Methods and Applications*, 20, 145–165, URL `http://www.degruyter.com/view/j/mcma.2014.20.issue-2/mcma-2013-0024/mcma-2013-0024.xml`.

Longstaff, Francis A and Eduardo S Schwartz (2001), "Valuing American Options by Simulation: A Simple Least-Squares Approach." *Review of Financial Studies*, 14, 113–147, URL `http://rfs.oxfordjournals.org/lookup/doi/10.1093/rfs/14.1.113`.

Tsitsiklis, John N. and Benjamin Van Roy (2001), "Regression methods for pricing complex American-style options." *IEEE Transactions on Neural Networks*, 12, 694–703.