

CONTROL VARIATE TECHNIQUE: A CONSTRUCTIVE APPROACH

Tarik Borogovac

Department of Electrical and Computer Engineering
& Center for Information and Systems Engineering
Boston University
8 Saint Mary Street
Boston, MA 02215, U.S.A.

Pirooz Vakili

Division of Systems Engineering
& Mechanical Engineering Department
Boston University
15 Saint Mary Street
Brookline, MA 02446, U.S.A.

ABSTRACT

The technique of control variates requires that the user identify a set of variates that are correlated with the estimation variable and whose means are known to the user. We relax the known mean requirement and instead assume the means are to be estimated. We argue that this strategy can be beneficial in parametric studies, analyze the properties of controlled estimators, and propose a class of generic and effective controls in a parametric estimation setting. We discuss the effectiveness of the estimators via analysis and simulation experiments.

1 INTRODUCTION

To achieve estimation efficiency, Variance Reduction Techniques (VRT) generally bring some *auxiliary information* to bear on the estimation problem. The Control Variate technique (CV) makes this use most explicit: Assume the simulation objective is to estimate the mean of a random variable Y ; The CV method relies on one or more auxiliary random variables called *controls* and utilizes information about these variables (their known means) to reduce the variance of the estimator for $E[Y]$. Hence, the CV method can be viewed as an approach to *extract and transfer information* included in controls where information is construed broadly.

Once the source of information, i.e., the set of controls, is identified, the mechanisms for optimal information extraction and transfer are well understood and analyzed (See, e.g., (Glasserman 2004), (Asmussen and Glynn 2007), (Nelson 1990), (Szechtman 2003), (Szechtman 2006), (Lavenberg and Welch 1981)). On the other hand, while there are some guidelines and common approaches for *identifying/discovering* controls, the identification/discovery process is fairly ad-hoc and its success depends critically on the ingenuity of the user.

An effective control, say X , needs to satisfy two requirements (to simplify the discussion we consider a scalar control):

1. X needs to be correlated with Y (carry some information about Y), and
2. $E[X]$ needs to be available to the user, i.e., known.

The critical barrier to finding effective controls is the second requirement, namely the requirement of a known mean $E[X]$. Users of the CV technique generally have two avenues to explore for satisfying the second requirement: (i) use some input (*internal*) random variable whose mean is given, or (ii) use an *external* control X whose mean can be analytically evaluated. Input random variables often do have some correlation with simulation output Y but this correlation is typically not very strong, limiting the effectiveness of such choices except in rare cases. For a generic simulation project with an output random variable Y it is often nontrivial to find a strongly correlated r.v. X whose mean can be evaluated analytically. This is mainly due to the limited reach of analytic methods.

A modification of the CV technique called Biased Control Variate (BCV) reduces the burden of requirement 2 by allowing for a good approximation of $E[X]$ when $E[X]$ cannot be evaluated analytically (see, e.g., (Schmeiser, Taaffe, and Wang 2001)). While BCV lowers the requirement barrier and expands the range of available choices for effective controls, it nonetheless limits its potential scope by implicitly assuming an analytic path to arriving at the approximate value (assumed to be a deterministic quantity). In (Schmeiser, Taaffe, and Wang 2001) it is pointed out that some practitioners use simulation to estimate $E[X]$ but the paper does not promote this choice and its analysis implicitly assumes an “analytic” approximation.

Allowing for estimation of $E[X]$ via simulation drastically enlarges the range of choices of control variates available to the user. On the other hand, this is a potentially costly proposal: To obtain a high quality estimate of $E[X]$

may require a large number of samples of X , and hence a high computational cost. In other words, a constructive and computational approach to obtaining control variates introduces the burden/challenge of intelligent management and justification of computational cost, an issue that requires careful consideration and analysis. From this perspective analytic evaluation of $E[X]$ or its “analytic” approximation can be viewed as extremely efficient methods of gathering information about $E[X]$, methods that, alas, are not available for many X ’s.

In this paper we pursue this direction of investigation. In other words, we assume that the information about X is to be computationally gathered and then used for estimation of Y^1 . We present two settings where the constructive approach can be quite effective: (1) In a parametric evaluation/optimization setting where the estimation variable Y depends on some model parameter (i.e., $Y = Y(\theta)$ for some $\theta \in \Theta$) and the simulation project may require simulations at many parameters. If one set of controls can be used to gain simulation efficiency at many parameters, the cost of estimating the mean of the controls can be viewed as an initial investment that accrues dividends in efficiency gains at each estimation instance. In some projects the dividends can easily justify the initial investment. (2) For estimation problems where there is a real-time constraint on estimating $E[Y]$, the estimation of means of controls can be viewed as a setup/off-line cost that can lead to significantly better estimates of $E[Y]$. Often this benefit more than justifies the setup cost.

The second thrust of the paper is to propose a *generic* class of control variates in the context of parametric estimation problems. The approach to deriving such control variates is independent of the specific estimation problem being considered and is not taxing on user ingenuity. The core idea of the approach is to use variables at one or more nominal parameters and possibly their sensitivities as controls for estimation at neighboring parameters. Namely $X_i^1 = Y(\theta_i)$, $X_i^2 = \frac{\partial Y}{\partial \theta}(\theta_i)$, $i = 1, \dots, k$ when $Y = Y(\theta)$ ((Zhao, Borogovac, and Vakili 2007a) and (Borogovac and Vakili 2008) include some experimental results from computational finance and computational physics domains respectively; (Borogovac and Vakili 2008) includes a detailed discussion). Under mild assumptions, these controls are highly correlated with Y and hence turn out to be very effective controls.

The rest of the paper is organized as follows. Preliminaries are given in Section 2. Section 3 provides an analysis of variance reduction for cases where the means of controls are estimated via simulation. Section 4 intro-

¹This is similar to bootstrapping (see, e.g., (Shao and Tu 1995)) where Monte Carlo sampling is used to estimate the variance of a statistics when an analytic formula for the variance is not available. Results from the analysis of bootstrap estimator turn out to be relevant to our discussions to follow.

duces parametric controls. Experimental results are given in Section 5 and we conclude in Section 6.

2 PRELIMINARIES

We begin with an overview of classical CV and Biased CV. This section also serves to establish some notation.

2.1 Classical Control Variate (CV)

Consider a probability space (Ω, \mathcal{F}, P) . Assume all random variables are defined on this space and all expectations are with respect to P . Let $L^2(\Omega, \mathcal{F}, P)$ ($L^2(\Omega)$ for simplicity) denote the set of random variables of finite second moment on this probability space. In what follows all random variables are assumed to belong to $L^2(\Omega)$. Let Y be a random variable whose mean

$$J = E[Y]$$

is unknown and is to be estimated. In addition, assume X_1, \dots, X_k are k control variates with known means $E[X_i] = \mu_i$. Let $\mathbf{X} = (X_1, \dots, X_k)$ and $\boldsymbol{\mu} = (\mu_1, \dots, \mu_k)$ (all vectors are assumed to be column vectors irrespective of how they are presented). The following class of (linear) controlled estimators of J is considered.

$$Z = Y - \beta'(\mathbf{X} - \boldsymbol{\mu})$$

where $\beta \in R^k$ and \prime denotes transpose. Z is the “corrected” or controlled estimator of J where the deviations of X_i ’s from their known means are used to “correct” samples of Y and draw them closer to its mean J and hence reduce estimator variance.

For fixed β , Z is an unbiased estimator of J . Assuming a nonsingular $\Sigma_{\mathbf{X}}$ the optimal β is given by

$$\beta^* = \Sigma_{\mathbf{X}}^{-1} \Sigma_{XY}$$

In this case

$$Var(Z) = (1 - R_{XY}^2) Var(Y)$$

where $R_{XY}^2 = \Sigma'_{XY} \Sigma_{\mathbf{X}}^{-1} \Sigma_{XY}$. Therefore, $(1 - R_{XY}^2)$ is the degree of variance reduction that can be achieved if the optimal controlled estimator is used compared to the crude MC estimator. It is worth noting that $L^2(\Omega)$ with the inner product

$$\langle X, Y \rangle = E[XY]$$

is a Hilbert space. The term $\beta^{*\prime}(\mathbf{X} - \boldsymbol{\mu})$ corresponds to the familiar orthogonal projection of Y on the linear subspace spanned by X_1, \dots, X_k in the Hilbert space setting. Z is the innovation term corresponding to the variation/information

in Y that is not explainable by the control variates (see, e.g., (Szechtman 2006)).

In general Σ_{XY} (and possibly Σ_X) is not known a priori. Therefore, the optimal β is not known a priori and needs to be estimated. The common estimator for β^* is

$$\hat{\beta} = S_X^{-1} S_{XY}$$

where S_X is the sample covariance matrix of X , and S_{XY} is the vector of sample covariances between Y and the components of X . The controlled estimator therefore is

$$\hat{Z} = Y - \hat{\beta}'(X - \mu).$$

For a more thorough discussion classical CV see, e.g., (Lavenberg and Welch 1981), (Nelson 1990), (Glasserman 2004).

2.2 Biased Control Variate (BCV)

BCV assumes means of the controls, i.e., μ_i ($i = 1, \dots, k$), are not known and instead good approximations of them, say, $\tilde{\mu}_i$ ($i = 1, \dots, k$), are available. To simplify the discussion assume $k = 1$. Let $\varepsilon = \tilde{\mu} - \mu$ denote the approximation error. For a fixed β , the BCV estimator is defined as

$$Z_B = Y - \beta(X - \tilde{\mu}).$$

Z_B is a biased estimator of μ with bias $\beta(\tilde{\mu} - \mu) = \beta\varepsilon$ and the mean square error of the estimator is

$$MSE(Z_B) = Var(Z_B) + (\beta\varepsilon)^2.$$

β may be selected to minimize $MSE(Z_B)$ or $Var(Z_B)$. In either case, as in classical CV, optimal β needs to be estimated from sampled data.

BCV expands the range of available controls at the cost of introducing some bias. Its use can be justified if the resulting reduction of variance is large enough to overcome the incurred bias. In cases where some information about the bias is available (say a bound on its magnitude) one can construct confidence intervals for J .

Consider the case where $\tilde{\mu}$ is exogenously available (exogenous to sampled data, say through some analytic approximation approach). In other words, assume obtaining more samples does not affect the approximation error ε . (Schmeiser, Taaffe, and Wang 2001) provides an insightful analysis of this case and show that when

$$n < \frac{\sigma_X^2}{\varepsilon^2}$$

(n is the number of i.i.d.samples) the variance reduction is significant enough to justify incurring the bias. Note

that this threshold does not depend on the correlation ρ_{XY} . In other words, as long as the bias is significantly smaller than the variance of the control, it pays to control the crude MC estimator. As n increases the variance of the control decreases while its bias remains unchanged. After the threshold provided above the disadvantages of the bias overwhelms the advantage of controlling the crude MC estimator.

3 ESTIMATED CONTROL VARIATE (ECV)

In this section we assume that the mean of the control is estimated via simulation. This approach is not new. It is known and utilized by some simulation practitioners (see, e.g., (Schmeiser, Taaffe, and Wang 2001)). We advocate lifting/promoting this approach from an ad-hoc method used by practitioners to an estimation approach worth general consideration and study and one that, as we argue, can be quite effective in some estimation problems.

We consider the following setting. To simplify the discussion we assume a single control X is used. Assume N samples (N “large”) are drawn from the probability space Ω according to probability measure P and $\hat{\mu}$, the estimate of μ , is calculated as the sample average

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N X(\omega_i).$$

Fixing $\Omega_{DB} = \{\omega_1, \dots, \omega_N\}$, $\hat{\mu}$ can be viewed as the expected value of the random variable X restricted to Ω_{DB} with respect to the uniform measure on this discrete probability space (all subsets are measurable). We refer to Ω_{DB} as the *database* and denote the restrictions of random variables to Ω_{DB} by subscript DB . For example, X_{DB} is the restriction of X to the database. We have

$$\hat{\mu} = E[X_{DB}]$$

where expectation is with respect to the uniform measure on Ω_{DB} . Alternatively, let P_{DB} denote the empirical measure associated with samples $\omega_1, \dots, \omega_N$ (Ω_{DB} for short). Then

$$\hat{\mu} = E_{DB}[X]$$

where E_{DB} denotes expectation with respect to P_{DB} .

3.1 Two implementations

In the following we consider two implementations of the ECV approach. In both cases we assume Ω_{DB} is generated and $\hat{\mu}$ is evaluated.

Figure 1 gives the schematic of the first implementation of ECV and shows how individual controlled samples are generated. Note that step 1 of the algorithm corresponds

-
1. Generate ω according to probability measure P_{DB}
 2. Evaluate $X_{DB} = X(\omega)$ and $Y_{DB} = Y(\omega)$
 3. Set

$$Z_{E1} = Y_{DB} + \beta(X_{DB} - \hat{\mu})$$

Figure 1: Implementation 1 of ECV

to sampling uniformly from the existing database Ω_{DB} (equivalently sampling from Ω according to the empirical measure P_{DB}). In this case, $E[X_{DB}] = \hat{\mu}$, i.e., the control variate is unbiased. However, $E[Y_{DB}]$ is in general different from the quantity we aim to estimate, i.e., $J = E[Y]$ and the controlled estimator Z_{E1} is a biased estimator of J with bias $\varepsilon_1 = E[Y] - E_{DB}[Y]$.

-
1. Generate ω according to probability measure P
 2. Evaluate $X = X(\omega)$ and $Y = Y(\omega)$
 3. Set

$$Z_{E2} = Y + \beta(X - \hat{\mu})$$

Figure 2: Implementation 2 of ECV

Figure 2 gives the schematic of the second implementation of ECV and again shows how individual controlled samples are generated. In this case step 1 of the algorithm corresponds to sampling from the original space Ω according to the original probability measure P . In this case, Y is an unbiased estimator of J , however $E[X - \hat{\mu}] \neq 0$, leading to a biased controlled estimator Z_{E2} with bias $\varepsilon_2 = \beta(E[X] - E_{DB}[X]) = \beta(\mu - \hat{\mu})$.

We make the following observations about the first implementation. Similar assertions can be made about the second.

- The first implementation of ECV can be viewed as the application of the classical control variate for random variables restricted to Ω_{DB} . The estimation variable is Y_{DB} and the control X_{DB} . In this case, the mean of the control, X_{DB} , is exactly evaluated via a “brute force” approach by taking a weighted average of all values that X_{DB} adopts.
- The controlled estimator is biased in both implementations. The bias in the case of the first implementation is due to the fact that controlled estimator is used to estimate the mean of an approximation, Y_{DB} , to the estimation variable Y . More

specifically,

$$\varepsilon_1 = E[Z_{E1}] - J = E_{DB}[Y] - E[Y]$$

- The bias ε_1 is zero on average (averaged over all databases), i.e. $E[\varepsilon_1] = 0$. Moreover, given the strong law of large numbers, for all databases, the bias is “small” for “large” N . In addition, we have

$$\sqrt{N} \frac{\varepsilon_1}{\sigma_{Y_{DB}}} \Rightarrow N(0, 1)$$

Therefore, the bias is of order $1/\sqrt{N}$ and it can be reduced by increasing N , i.e. increasing the size of the database Ω_{DB} . From the above a probabilistic assessment of the magnitude of the bias is readily available.

$$P(|\varepsilon_1| < c \frac{\sigma_Y}{\sqrt{N}}) \approx 1 - 2\Phi(c)$$

where Φ is the cumulative distribution of the standard normal and we consider the approximation $\sigma_{Y_{DB}} \approx \sigma_Y$.

- Assume n samples are drawn from Ω_{DB} and let $\bar{Z}_{E1}(n)$ be the sample average of n controlled samples. Then

$$E[(\bar{Z}_{E1}(n) - J)^2] = \frac{\sigma_Y^2}{n} (1 - \rho^2) + \frac{\sigma_Y^2}{N}$$

Note that the above expectation is taken over all possible Ω_{DB} and the identity reflects the average performance of the first implementation of ECV. At the cost of introducing a squared bias term that on average is about σ_Y^2/N one can potentially arrive at control variates with higher correlation ρ . The above identity provides the basis for the assessment of the tradeoff.

- Given that a probabilistic assessment of the bias term can be obtained, we can construct confidence intervals for J .

4 A GENERIC CLASS OF CONTROLS

As mentioned earlier, a main motivation for using the ECV approach is to enlarge the range of effective controls that can be utilized. Furthermore, given that there are no barriers to approximating the mean of the controls except consideration of computational cost, one hopes to be able to develop a generic approach for selecting effective control variates. We describe an attempt in this direction in the setting of parametric simulation/estimation as described below.

Assume the “estimation variable” $Y = Y(\theta)$ depends on some parameter $\theta \in \Theta$. Consider the following two classes of estimation problems:

- Problem A. The estimation problem needs to be repeatedly solved for different values of $\theta \in \Theta$.
- Problem B. The estimation problem needs to be solved for some values of $\theta \in \Theta$. These values are not known in advance. Once a parameter value, say θ becomes known, there are real-time computational/time constraints for obtaining an estimate of $J(\theta) = E[Y(\theta)]$.

For fixed $\omega \in \Omega$, $Y(\omega, \theta)$ is a function of θ . Therefore, assuming the same random input ω is used to generate samples of $Y(\theta)$ at different θ , $Y(\theta)$ can be viewed as a random real-valued function on Θ . Under some mild continuity assumptions on the sample functions (continuity in θ) one expects $Y(\theta_1)$ $Y(\theta_2)$ to be correlated when θ_1 and θ_2 are close. This motivates using $X_i = Y(\theta_i)$ for some select values of $\theta_i \in \Theta$, $i = 1, \dots, k$ as controls.

If, in addition, the sample functions are differentiable with respect to θ , it is possible to efficiently estimate sample path derivatives, and these estimates are unbiased, one may include $X_{i1} = \frac{\partial Y}{\partial \theta}(\theta_i)$ as additional set of controls.

We make the following observations

- Let $Y = Y(\theta_0)$. It is possible to show that under mild assumptions, one can find $X = Y(\theta)$ correlated with Y and with correlation as close to 1 as desired by letting θ approach θ_0 .
- The above approach is most effective in cases where correlation relatively persists for θ reasonably far from θ_0 . The experimental results section of the paper includes some examples where the effectiveness of such controls is examined.
- We expect that the selections of the number of controls and their location to be problem dependent. Further study is needed to come up with some guidelines for this selection problem.
- Once the controls are selected, as in the previous section, two implementations are possible. In one case the database Ω_{DB} (or enough information about it to allow resampling) is maintained and all sampling correspond to resampling from the database. In the second, once means of controls are estimated, the database is discarded.

We now give a brief discussion of the computational cost of the approach. Generating a database and/or estimating the means of the controls correspond to an initial “setup” cost. To assess the cost, let C be the computational cost of generating a sample of $Y(\theta_i)$, i.e. generating $Y(\omega; \theta_i)$, $i = 1, \dots, k$. To simplify, we assume this cost is the same for all θ_i and ω .

Then the set-up cost of generating the database and obtaining averages of the controls is approximately $N \times k \times C$. Let $VRR(\theta)$ denote the variance reduction ratio at θ , i.e., the ratio of the variance of an uncontrolled sample and that of a controlled sample at θ . Then, the statistical error of a controlled estimator based on n samples is the same as that of $n \times VRR(\theta)$ samples of an uncontrolled estimator. (We assume that for large N , the induced biased of the approach is small and can be disregarded). Thus, the ratios of the computational costs of the two estimators (to arrive at the same statistical accuracy) is $(n \times VRR(\theta) \times C)/(n \times C) = VRR(\theta)$. Therefore, $VRR(\cdot)$ can serve as a measure of benefit of the approach.

For problem A above, if the total number of instances is sufficiently large, then the set-up cost becomes negligible and the approach is worthwhile. For problem B, the setup cost can be viewed as an off-line cost enabling potentially significant efficiency gains in the critical task of real-time estimation. Typically, the “cost” of delay in such real-time estimation problems is high and not merely computational, justifying a large computational effort off-line. We expect that the approach, if applicable to a real-time setting, is almost always worthwhile.

5 DATABASE MONTE CARLO (DBMC)

The method described in this paper, which specifically utilizes control variates, is a special case of a broader estimation strategy, named DataBase Monte Carlo (DBMC). The overarching idea of DBMC is that computational effort invested in one estimation problem may lead to more precise or computationally efficient estimators for related problems. As discussed above, and demonstrated in our numerical examples below, the control variate method can be a convenient and effective way to translate information from one problem to the other, but it is certainly not the only way, or necessarily the most advantageous. For example, other well known variance reduction methods are used for that purpose in (Zhao, Zhou, and Vakili 2006, Zhao, Borogovac, and Vakili 2007b, Zhao, Borogovac, and Vakili 2007a).

6 NUMERICAL EXAMPLES

6.1 Simple functions

To illustrate the implementation and give results, first we looked at estimation of the mean of random variables based on simple functions, which we denote $Y(\theta) = g(\theta, U)$, and which have the derivative $Y'(\theta) = \frac{dg(\theta, U)}{d\theta}$. As described above, we use a nominal parameter value θ_0 , where we “collect” the control variate samples $Y(\theta_0, U_i)$, $i = 1, \dots, N$, for use in estimation of $E[Y(\theta)] = \int_0^1 g(\theta, u) du$ on an interval of the parameter space $\Theta = [\theta_{\min}, \theta_{\max}]$.

Figures 3 and 4 show the squared correlation of the control variate at θ_0 with the estimation variables over a range of θ for estimating the means $E[g(\theta, U)]$ for two simple functions $g(\theta, U) = e^{\theta U}$ and $g(\theta, U) = \sin(\theta U)$, respectively. Squared correlations are used for presentation purposes instead of the equivalent variance reduction ratios (recall that for control variates $VRR = (1 - R^2)^{-1}$, thus R^2 values close to 1 are desirable). In both cases, the correlations obtained by using $Y(\theta_0)$ as the control imply significant variance reduction for a subset of the parameter space Θ (solid curve), but can still be greatly improved by using the derivative $Y'(\theta_0) = \frac{dg(\theta_0, U)}{d\theta}$ as the second control (dashed curve). For convenience, corresponding VRR values for estimation of $E[e^{\theta U}]$ at a few selected θ values are given in Table 1.

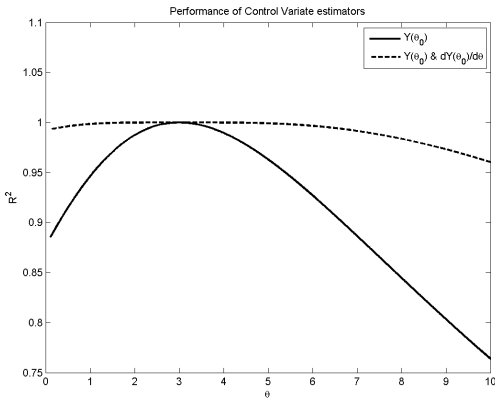


Figure 3: The squared correlation measuring the utility of control variates obtained by DBMC for estimation $E[Y(\theta)]$ for $g(\theta, U) = e^{\theta U}$. The nominal $\theta_0 = 3$

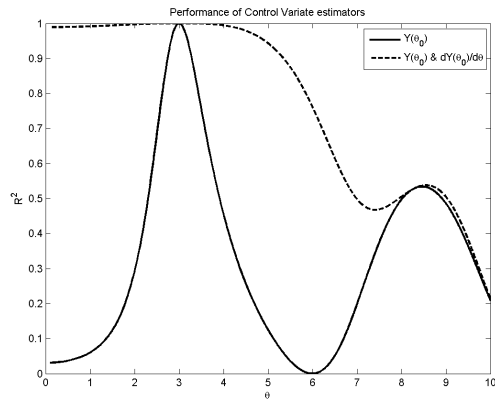


Figure 4: The squared correlation measuring the utility of control variates obtained by DBMC for estimation $E[Y(\theta)]$ for $g(\theta, U) = \sin(\theta U)$. The nominal $\theta_0 = 3$

Table 1: The values of the VRR statistic for selected points, corresponding to squared correlations in Figure 3.

Control \ θ	1	2	3.2	6
$Y(\theta_0)$	18	79	2.2E3	14
$Y(\theta_0) \& Y'(\theta_0)$	660	1.17E4	8.8E6	320

6.2 Asian option

We also illustrate how the algorithm can be used to estimate the price of financial options by describing their application to the valuation of an arithmetic Asian option on a Geometric Brownian model. There is no closed-form solution for the price of this option and simulation is a solution method of choice.

Specifically, the stock price is modeled as a GBM:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

where W_t is a Brownian motion. Given a discretized path, $S_0 = \{S_{t_0}, S_{t_1}, S_{t_2}, \dots, S_{t_m} = S_T\}$, the present value of the payoff of the Asian call option is given by

$$C_T = e^{-\mu T} [\bar{S} - K]^+, \quad \bar{S} = \frac{1}{m} \sum_{i=1}^m S_{t_i}$$

We are interested in estimation of the option price (of the discretized process), given by $E[C_T]$ at a range of drift parameter μ and volatility parameter σ . We can also make use of the pathwise derivatives

$$\frac{dC_T}{d\mu} = e^{-\mu T} I\{\bar{S} > K\} \left\{ \frac{1}{m} \sum_{i=1}^m \Delta t_i S_{t_i} - T[\bar{S} - K]^+ \right\}$$

where $\Delta t_i = t_i - t_{i-1}$, and

$$\frac{dC_T}{d\sigma} = e^{-\mu T} I\{\bar{S} > K\} \frac{1}{m} \sum_{i=1}^m \frac{dS_{t_i}}{d\sigma}$$

where, for all t :

$$\frac{dS_t}{d\sigma} = \frac{S_t}{\sigma} \left(\log \left(\frac{S_t}{S_0} \right) - \left(\mu + \frac{\sigma^2}{2} \right) t \right).$$

Note that we do not require unbiasedness of the path derivative in the sense that $E[\frac{dC_T}{d\mu}] = \frac{dE[C_T]}{d\mu}$ or $E[\frac{dC_T}{d\sigma}] = \frac{dE[C_T]}{d\sigma}$ which would be necessary in problems where the objective is estimation of price sensitivities *Rho* and *Vega* (although that unbiasedness condition is satisfied for the

Asian option by the path derivatives above (Glasserman 2004)). For use as a control variate in the estimation of the price, we only need the values of the path derivative for each sampled path, and their mean values $E[\frac{dC_T}{d\mu}]$ and $E[\frac{dC_T}{d\sigma}]$, which we estimate precisely by utilizing all the paths in the database.

The following parameters were considered: $S_0 = 100$, $K = 90$, $T = .25$, $m = 60$. The nominal values of the drift and volatility parameters were $\mu_0 = 0.1$, and $\sigma_0 = 0.3$, at which a database of 100,000 GBM paths was generated, and on each of those paths, we evaluated the Asian option price C_T , and the pathwise derivatives $\frac{dC_T}{d\mu}$ and $\frac{dC_T}{d\sigma}$. Those results were saved, as well as a random number generator seed for each of the paths, which were then used to estimate at other parameter values via the CV procedure described above. Figures 5 and 6 show the VRR results for estimators using $Y(\theta_0)$ (solid curve) as well as $Y(\theta_0)$ together with $Y'(\theta_0)$ (dashed curve) as control variates. For selected parameter values marked in the plots, the values are shown in Tables 2 and 3.

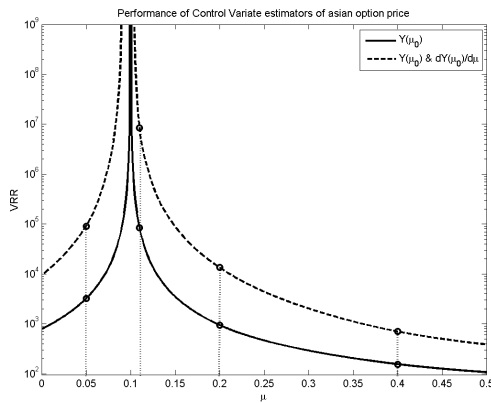


Figure 5: The VRR statistic measuring the efficiency of estimation of Asian option value over a range of drift rates μ when using a CV-DBMC estimator. Using the pathwise Rho , calculated at the nominal value μ_0 , as an additional control adds efficiency beyond what is provided by only using the nominal path option value $Y(\mu_0)$.

Table 2: The values of the VRR statistic for selected points, as shown in Figure 5.

Control \ μ	0.05	0.11	0.2	0.4
$Y(\mu_0)$	3.2E3	8.4E4	9.5E2	1.6E2
$Y(\mu_0) \& Y'(\mu_0)$	9.1E4	8.5E6	1.3E4	7.0E2

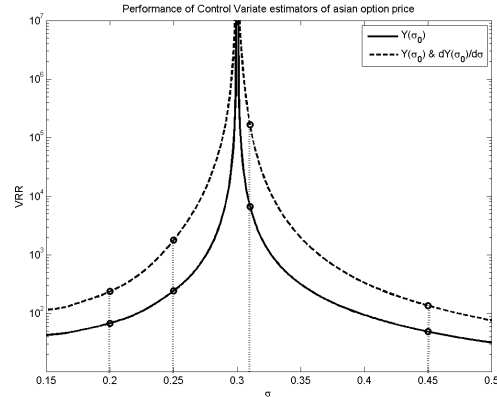


Figure 6: The VRR statistic measuring the efficiency of estimation of Asian option value over a range of volatilities σ when using a CV-DBMC estimator. Using the pathwise $Vega$, calculated at the nominal value σ_0 , as an additional control adds efficiency beyond what is provided by only using the nominal path option value $Y(\sigma_0)$.

Table 3: The values of the VRR statistic for selected points, as shown in Figure 6.

Control \ σ	0.2	0.25	0.31	0.45
$Y(\sigma_0)$	69	2.4E2	6.7E3	49
$Y(\sigma_0) \& Y'(\sigma_0)$	2.4E2	1.8E3	1.7E5	1.4E2

7 CONCLUSIONS

We have implemented a strategy for obtaining effective control variates in parametric studies. The strategy is quite problem independent and generic, as it can wrap around many different types of simulations. The controls obtained are usually quite effective in some neighborhood of the nominal parameter, as shown in our numerical examples, although the size of that neighborhood depends on the functional form of the estimation problem with respect to the parameter. Addressing that issue is an interesting problem for future research.

Another demonstrated advantage of the approach is the ability to straightforwardly obtain multiple control variates. In our numerical examples we found the pathwise derivatives to be quite effective as additional controls.

The up-front computational cost of generating the necessary database can be significant, and we have noted that it is only justified if either sufficiently many estimation problem instances benefit from that database, or if it is used in a real-time estimation setting. Additionally, the controls obtained incur some bias, but the magnitude of

that bias is under the practitioner's control when deciding the size of the database, and a probabilistic bound on the bias is straightforwardly assessed. The cost-benefit analysis provided illustrates the tradeoffs involved.

ACKNOWLEDGMENTS

Research supported in part by the National Science Foundation grants CMMI-0620965 and DGE-0221680.

REFERENCES

- Asmussen, S., and P. Glynn. 2007. *Stochastic simulation: Algorithms and analysis*. Springer.
- Borogovac, T., and P. Vakili. 2008. Database monte carlo approach to effective control variates. Technical report, Boston University College of Engineering.
- Glasserman, P. 2004. *Monte carlo methods in financial engineering*. Springer-Verlag New York, Inc.
- Lavenberg, S. S., and P. D. Welch. 1981. A perspective on the use of control variables to increase the efficiency of monte carlo simulations. *Management Science* 27 (3): 322–335.
- Nelson, B. L. 1990. Control variate remedies. *Operations Research* 38:974–992.
- Schmeiser, B., M. R. Taaffe, and J. Wang. 2001. Biased control-variate estimation. *IIE Transactions* 33:219–228.
- Shao, J., and D. Tu. 1995. *The jackknife and bootstrap*. Springer.
- Szechtman, R. 2003. Control variates techniques for monte carlo simulation. In *Proceedings of the 2003 Winter Simulation Conference*, ed. S. Schick, P. Sánchez, D. Ferrin, and D. J. Morrice, 144–149.
- Szechtman, R. 2006. A hilbert space approach to variance reduction. In *Handbook in OR and MS*, ed. S. G. Henderson and B. L. Nelson, Volume 13, Chapter 10, 259–289. Elsevier B.V.
- Zhao, G., T. Borogovac, and P. Vakili. 2007a. Efficient estimation of option price and price sensitivities via structured database monte carlo (SDMC). In *Proceedings of the 2007 Winter Simulation Conference*, ed. S. G. Henderson, B. Biller, M. H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 984–991.
- Zhao, G., T. Borogovac, and P. Vakili. 2007b. Structured database monte carlo: A new strategy for efficient simulation. Technical report, Boston University College of Engineering.
- Zhao, G., Y. Zhou, and P. Vakili. 2006. A new efficient simulation strategy for pricing path – dependent options. In *Proceedings of the 2006 Winter Simulation Conference*, ed. L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 703–710.

AUTHOR BIOGRAPHIES

TARIK BOROGOVAC is a Ph.D. student of Systems Engineering at Boston University. His current research interests include efficient Monte Carlo simulation in the areas of finance, imaging, and physical and biochemical systems. His e-mail address is <tarikb@bu.edu>.

PIROOZ VAKILI is an Associate Professor in the Division of Systems Engineering and the Department of Mechanical Engineering at Boston University. His research interests include Monte Carlo simulation, optimization, computational finance, and bioinformatics. His email address is <vakili@bu.edu>.