


## Article

# A Generalized Weighted Monte Carlo Calibration Method for Derivative Pricing

Hilmar Gudmundsson<sup>1,2</sup> and David Vyncke<sup>1,\*</sup> 

<sup>1</sup> Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281-S9, B-9000 Ghent, Belgium

<sup>2</sup> Verna, Ármúli 13, 108 Reykjavík, Iceland; hilmar@verna.is

\* Correspondence: david.vyncke@ugent.be

**Abstract:** The weighted Monte Carlo method is an elegant technique to calibrate asset pricing models to market prices. Unfortunately, the accuracy can drop quite quickly for out-of-sample options as one moves away from the strike range and maturity range of the benchmark options. To improve the accuracy, we propose a generalized version of the weighted Monte Carlo calibration method with two distinguishing features. First, we use a probability distortion scheme to produce a non-uniform prior distribution for the simulated paths. Second, we assign multiple weights per path to fit with the different maturities present in the set of benchmark options. Our tests on S&P500 options data show that the new calibration method proposed here produces a significantly better out-of-sample fit than the original method for two commonly used asset pricing models.

**Keywords:** option pricing; model calibration; weighted Monte Carlo



**Citation:** Gudmundsson, H.; Vyncke, D. A Generalized Weighted Monte Carlo Calibration Method for Derivative Pricing. *Mathematics* **2021**, *9*, 739. <https://doi.org/10.3390/math9070739>

Academic Editor: Elisa Alòs

Received: 14 February 2021

Accepted: 24 March 2021

Published: 29 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

One key metric used to gauge the success of an arbitrage option pricing model is how well it can reproduce the observed prices of liquid options in the market. In the context of the Black–Scholes framework of Black and Scholes [1], an equivalent question is how closely the implied volatility surface of the model can be made to resemble the volatility surface implied by the liquid options in the market.

Given the well-known limitations [2] of the Black–Scholes model, such as the assumption of constant volatility in the geometric Brownian motion, numerous generalizations of the model have been introduced with the aim of generating a better fit to stylized market features. Well-known modifications include everything from making the volatility a stochastic process [3] and/or a function of the underlying [4,5] to adding jump components to the underlying and volatility [6–8].

The downside of working with these advanced models includes a higher computational burden for pricing and an increased number of parameters that need to be estimated [9,10]. In particular, if an option pricing model implies market incompleteness, e.g., by including unobservable sources of randomness, such as jumps or stochastic volatility, more than one combination of parameter values can produce the prices we see, which can further complicate the calibration process. In addition, to use the model to price derivatives, we must make parametric assumptions on the market price of risk, i.e., implicitly make assumptions about the preferences of the market participants [11] (pp. 209–229). Preferences can also be viewed as being implicit in the calibration of option pricing models to financial data in general; without a perfect model, the calibration process is inherently a decision-theoretic problem, as different loss functions can lead to different economic consequences for the model user. We refer to the work of Friedman et al. [12] for an in-depth discussion of this point.

Instead of working with advanced models, Avellaneda et al. [13] proposed to “correct” a simple model in such a way that it reproduces the prices of the derivatives used in the

model calibration. Their weighted Monte Carlo (WMC) method consists of simulating a set of price paths using the arbitrage model that is to be calibrated and calculating a new (risk neutral) probability measure for this set of paths that reproduces the observed market prices of benchmark options exactly, or almost exactly, as in the case of a least squares approach.

As we tend to have more paths than benchmark options in a simulation, such a measure is not uniquely defined in general. This problem is solved in the original WMC method by selecting the measure closest to the (uniform) prior measure in terms of relative entropy, also referred to as Kullback–Leibler divergence [14]. As Kullback and Leibler [14] pointed out, relative entropy is a justified choice of a statistical divergence due to its mathematical tractability and the fact that it is invariant under change of variables. Furthermore, when we are faced with the problem of picking the “best” probability distribution from a set of distributions that all fit with a given set of observed data, the most parsimonious choice from an information-theoretic perspective is the one that exhibits minimal relative entropy [15,16].

While relative entropy is a well established concept within information theory, it does not by itself offer any particular economic intuition in the context of the WMC method. The implication of this is that there is no obvious way to discriminate between relative entropy and any other kind of divergence when choosing how to weight the sample paths in a manner that is economically justified. However, as we discuss in greater detail in the following sections, the entropy minimization problem of the WMC method is mathematically equivalent to a portfolio choice problem for an investor with expected exponential utility. By considering the utility version of the WMC, we open up the possibility of exploiting the theoretically and empirically mature field of consumption based asset pricing [17,18] to develop refinements of the WMC method.

As we explain in greater detail in the following sections, the WMC calibration method can in theory reweight the paths simulated from the model to be calibrated in such a way that they can reproduce the market prices of the benchmark options used in the calibration procedure exactly. However, as we observe in our numerical tests for two popular option pricing models (the Black–Scholes model [1] and the stochastic volatility model of Heston [3]), the accuracy drops quite quickly for out-of-sample options as we move away from the strike range and maturity range of the benchmark options.

The contribution of this paper consists of formulating a more general version of the WMC which in our numerical tests produces a far better fit to the whole range of options available on the underlying asset than the original WMC. It achieves this by first splitting the paths into segments by the maturities present in the set of benchmark options, and then applying a probability distortion transformation to the prior distribution associated with these path segments.

This probability distortion transformation is inspired by the work of Tversky and Kahneman [19] on Cumulative Prospect Theory (CPT). On a theoretical level, CPT has been shown to produce potential resolutions to a wide array of empirical puzzles in finance and economics [20–25]. As compared to expected utility, two features characterize CPT preferences. The first feature is a utility function over deterministic outcomes that is concave over gains and convex over losses with respect to a given reference point of wealth. The second feature is a distortion function that is applied to the cumulative distribution function of the physical measure. Although the probability distortion generally leads to a non-additive expectation operator, the formulation we propose in this paper maintains the additivity of the expectation operator.

The remainder of this paper is as follows. In Section 2, we give a general formulation of the weighted Monte Carlo method. In Section 3, we discuss the relationship between entropy minimization and utility maximization and give an alternative formulation of the entropy minimization problem as a portfolio choice problem. In Section 4, we propose a weighted Monte Carlo method that incorporates multiple weights per path and rare event probability distortion. In Section 5, we give the numerical results for the different methods.

## 2. An Overview of the Weighted Monte Carlo Method

We begin by briefly describing the weighted Monte Carlo method in general terms. Our discussion follows the same reasoning as was presented in a more comprehensive format by Avellaneda et al. [13], with the exception that we do not assume that the prior distribution is uniform. Thus, the original WMC method is a special case of what is presented in this section.

Given a filtered probability space  $(\Omega, \mathcal{F}, F, \mathbb{P})$ , a finite time horizon  $T$ , and an adapted price process  $S = \{S_t\}_{t \in [0, T]}$ , we can view the set of  $N$  paths produced by a Monte Carlo simulation of  $S_t$  as a discrete approximation to the distribution of  $S$  at time  $t$ . The general idea behind the weighted Monte Carlo approach is to reweight the sampled paths in such a way that the new distribution is as statistically close as possible to the original one, while at the same time reproducing the observed market prices at time 0 for a set of contingent claims on  $S$ . Here, the notion of statistical distance is taken to be the Kullback–Leibler divergence. However, other types of divergences have been studied, see [26]. The Kullback–Leibler divergence for two discrete probability measures,  $\mathbb{P}$  and  $\mathbb{Q}$ , is given by

$$D_{KL}(\mathbb{Q}|\mathbb{P}) = \sum_i q_i \ln\left(\frac{q_i}{p_i}\right). \tag{1}$$

We refer to the no-arbitrage model to be calibrated using the weighted Monte Carlo method as the initial model. Assume we have a set of  $K$  benchmark options on  $S$ , which we want to use to reweight the simulated paths of the model. Any type of option will do as long as its payoff along a given sample path is completely determined by that path, which excludes, for example, American style options. Assume further that no option in the set can be replicated by a portfolio of the other remaining options. We begin by simulating  $N$  paths from  $S$  and calculating the stochastic payoffs  $(G_1, \dots, G_K)$ . This gives the payoff matrix

$$G = \begin{pmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,K} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ g_{N,1} & g_{N,2} & \cdots & g_{N,K} \end{pmatrix}, \tag{2}$$

where  $g_{ik}$  is the payoff from option  $k$  when path  $i$  is realized.

Assuming the simulated paths are initially assigned prior weights given by  $\mathbf{p} = (p_1, \dots, p_N)$ , the new weights  $\mathbf{q} = (q_1, \dots, q_N)$  in the exact-fit version are calculated as the solution to

$$\begin{aligned} \min_{\mathbf{q}} D(\mathbf{q}|\mathbf{p}) &= \min_{\mathbf{q}} \left\{ \sum_{i=1}^N q_i \ln\left(\frac{q_i}{p_i}\right) \right\} \\ \text{s.t. } \pi_k &= \sum_{i=1}^N q_i g_{ik}, \quad k = 1, \dots, K \end{aligned} \tag{3}$$

where  $\pi_k$  is the price of benchmark option  $k$ . Note that throughout our discussion we use  $\mathbb{Q}$  and  $\mathbf{q}$  (and  $\mathbb{P}$  and  $\mathbf{p}$ ) interchangeably but write the former when we want to emphasize it as a probability measure and the latter when we want to emphasize it as a set of path weights.

Using the Lagrange multiplier approach, this can be rewritten as the dual problem

$$\min_{\lambda} \max_{\mathbf{q}} \left\{ -D(\mathbf{q}|\mathbf{p}) + \sum_{k=1}^K \lambda_k \left( \sum_{i=1}^N q_i g_{ik} - \pi_k \right) \right\}. \tag{4}$$

Looking at the first order conditions for the inner problem, we see that a solution is given by

$$q_i = \frac{p_i}{Z(\lambda)} \exp\left(\sum_{k=1}^K g_{ik} \lambda_k\right), \tag{5}$$

where

$$Z(\lambda) = \sum_{m=1}^N p_m \exp\left(\sum_{k=1}^K g_{mk}\lambda_k\right)$$

is a normalization factor to ensure the distribution sums to 1. Using (5), we can simplify (4) to get

$$\min_{\lambda} \left\{ \ln(Z(\lambda)) - \sum_{k=1}^K \pi_k \lambda_k \right\}. \tag{6}$$

This problem requires numerical methods to solve. Given that the objective function is smooth and convex, the method of choice is a gradient-based minimization procedure such as the BFGS algorithm [27] which requires the computation of the partial derivatives of (6). Defining

$$W(\lambda) = \ln(Z(\lambda)) - \sum_{k=1}^K \pi_k \lambda_k \tag{7}$$

they are given by

$$\begin{aligned} \frac{\partial W}{\partial \lambda_j} &= \frac{1}{Z(\lambda)} \sum_{i=1}^N g_{ij} p_i \exp\left(\sum_{k=1}^K g_{ik}\lambda_k\right) - \pi_j \\ &= \mathbb{E}^Q[G_j] - \pi_j \end{aligned} \tag{8}$$

for  $j = 1, \dots, K$ . Taking second derivatives, we get

$$\frac{\partial^2 W}{\partial \lambda_i \partial \lambda_j} = \text{Cov}(G_i, G_j) \tag{9}$$

We recall that a covariance matrix is non-negative definite by definition. Since we assume that (2) is of full rank (from our earlier assumption that the set of options contains no redundant options), we have that the covariance matrix is in fact positive definite. This implies that the function is strictly convex, which in turn implies that any solution to (8) corresponds to a global minimum.

Finally, although a unique solution should always exist for (3) given an arbitrage free market, the presence of asynchronous and noisy data can lead to problems for the optimization procedure, with one instrument being bought or sold in quantities much larger than the rest. A remedy to this problem is adding a regularization term to the objective function. The term we use in our tests is a simple quadratic term,

$$\chi_w^2 = \frac{1}{2} \sum_{k=1}^K w_k \lambda_k^2, \tag{10}$$

where  $w_1, \dots, w_K$  are penalization weights that determine the influence of option  $k$  on the calibration. The resulting optimization problem in (6) is given by

$$\min_{\lambda} \left\{ \ln(Z(\lambda)) - \sum_{k=1}^K \pi_k \lambda_k + \chi_w^2 \right\}. \tag{11}$$

### 3. The Weighted Monte Carlo Method as a Utility Maximization Problem

Using the observed price information of assets such as stocks and options in a financial market to derive (or approximate) the stochastic discount factor of that market is a central task in the field of asset pricing (for a standard reference, see, e.g., [28]). Within the field of financial economics, this is commonly done in the setting of consumption based pricing, where the market is treated as a single representative investor with a given utility function over uncertain future cash flows. The stochastic discount factor in this setting

is proportional to the marginal utility of the representative investor over her optimal consumption in each possible future state. Given a set of contingent claims in the market, this optimal consumption process can be calculated by solving for the optimal portfolio of the investor.

In more concrete terms, consider the following one period model. Here, ‘one period’ refers to the investor only choosing a portfolio at time 0. This does not restrict the assets from evolving continuously from time 0 to time  $T$  with realized payoffs in between. We assume the investor maximizes expected utility, with investment horizon  $T$ . The assets available to the investor form an arbitrage-free market of total initial wealth  $W_0$  with an associated probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let the investor’s utility function over deterministic outcomes realized at time  $T$  be given by a continuously differentiable function  $u : \mathbb{R} \rightarrow \mathbb{R}$  such that  $u'(x) > 0$  and  $u''(x) < 0$  for all  $x \in \mathbb{R}$ , and let  $\mathcal{A}(W_0)$  denote the set of all contingent claims in the market which can be financed with initial wealth  $W_0$ . If we denote with  $X_T^*$  the optimal solution to the portfolio selection problem

$$\sup_{X_T \in \mathcal{A}(W_0)} \mathbb{E}^{\mathbb{P}}[u(X_T)], \tag{12}$$

there exists  $\beta \in \mathbb{R}^+$  such that for every contingent claim  $X_T$  in the market we have that the market price of  $X_T$  is given by  $\pi_X = \mathbb{E}^{\mathbb{P}}[\beta u'(X_T^*) X_T]$ . We can write the expectation in more compact form through a change of measure using the Radon–Nikodym derivative  $\frac{d\mathbb{Q}}{d\mathbb{P}} = u'(X_T^*)$  to obtain  $\pi_X = \mathbb{E}^{\mathbb{Q}}[\beta X_T]$ .

In mathematical terms, the correspondence between this type of utility maximization and entropy minimization is well known [29]. It is therefore reasonable to ask precisely how this relationship enters in the weighted Monte Carlo setting. As explained in [13], the Arrow–Debreu prices that correspond to the measure computed from (3) coincide with the marginal utilities for consumption obtained by solving (12) when  $u$  is the exponential utility function.

More specifically, solving for the Lagrange multipliers  $\lambda_k$  in (4) is equivalent to finding the optimal portfolio weights for a utility maximizer with exponential utility [30], where we can think of the benchmark instruments as the assets in the market, and the set of sample paths generated by the initial model as the state space of the market. The relationship between the two is more precisely as follows: the optimal portfolio weights given by  $\phi = (\phi_1, \dots, \phi_K)$  for an investor with a utility function given by

$$u(\phi) = - \sum_{i=1}^N p_i \exp \left( - \sum_{k=1}^K \phi_k (g_{ik} - \pi_k) \right) \tag{13}$$

are related to the optimal Lagrange multipliers  $\lambda_1, \dots, \lambda_K$  by  $\phi_k = -\lambda_k$  for  $k = 1, \dots, K$ . Consequently, if we assume that the utility maximizer is a representative investor, then the path weighting through entropy minimization is mathematically equivalent to deriving the stochastic discount factor in this market through solving for the optimal portfolio of the representative investor.

Before elaborating further on how the weighted Monte Carlo method fits with the standard consumption-based asset pricing approach, we need to introduce some additional terminology. Given an underlying state space  $\Omega$ , and a  $\sigma$ -algebra  $\mathcal{F}$  defined on  $\Omega$ , we refer to the probability measure that describes the true probability of the different events in  $\mathcal{F}$  as the *objective* measure. In contrast, we refer to any probability measure that is absolutely continuous with respect to the objective measure, but absorbs to some extent risk premiums that exist in the market, as a *subjective* measure. An example of a subjective measure would be a risk neutral measure, i.e., a probability measure  $\mathbb{Q}$  such that the market price of any contingent claim in the market is the discounted expected value of the claim with respect to  $\mathbb{Q}$ . Lastly, we use ‘pricing measure’ and ‘risk neutral measure’ interchangeably.

The standard way to formulate a consumption-based asset pricing model is to assume that the investor observes the objective measure  $\mathbb{P}$  associated with the state space.

This means the expected utility is calculated under the  $\mathbb{P}$  measure. On the other hand, option pricing models are typically calibrated against existing option contracts, making the probability measure corresponding to the sample paths of the initial model subjective, as opposed to objective.

However, this apparent discrepancy is dissolved with the realization that we are effectively deriving a “decomposed” stochastic discount factor. To clarify, let  $\mathbb{L}$  be the subjective probability measure corresponding to the uniform weighting of the Monte Carlo simulation paths generated by the initial model (i.e.,  $\mathbb{L}$  is the subjective measure under which the sample path approximation to the initial model is given). Let  $\mathbb{P}$  denote the objective probability measure corresponding to these paths. If our model is arbitrage free, we expect that  $\mathbb{L} \ll \mathbb{P}$ , i.e., that  $\mathbb{L}$  is absolutely continuous with respect to  $\mathbb{P}$ . Next, let  $\mathbb{Q}$  be the subjective probability measure corresponding to the weighting of the sample paths computed by the WMC method. Again, the absence of arbitrage means we must have  $\mathbb{Q} \ll \mathbb{L}$ . The Radon–Nikodym derivative  $\frac{d\mathbb{Q}}{d\mathbb{L}}$  is proportional to the stochastic discount factor  $u'(X_T^*)$  we obtain from the utility maximization equivalent of the WMC entropy minimization. Given that  $\mathbb{Q} \ll \mathbb{L} \ll \mathbb{P}$ , then, by the measure-theoretic chain rule, we have that

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{d\mathbb{Q}}{d\mathbb{L}} \frac{d\mathbb{L}}{d\mathbb{P}} - \mathbb{P} \text{ a.s.} \tag{14}$$

In other words, changing the measure from  $\mathbb{P}$  to  $\mathbb{L}$  and then from  $\mathbb{L}$  to  $\mathbb{Q}$  is the same (a.s.) as changing it straight from  $\mathbb{P}$  to  $\mathbb{Q}$ . Thus, whether we derive the risk neutral pricing distribution  $\mathbb{Q}$  straight from  $\mathbb{P}$ , as in the standard representative investor pricing approach, or by first deriving  $\mathbb{L}$  from  $\mathbb{P}$  and then  $\mathbb{Q}$  from  $\mathbb{L}$ , as in the WMC approach, we end up with the same results.

With the theoretical justification out of the way, we now give the utility maximization equivalent of the formulation presented in Section 2. Let  $u(\cdot)$  be defined as in (12). The state space  $\Omega$  consists of the paths we simulate from the initial model, and  $\mathbb{P}$  is whatever weighting we attach to the paths to represent the prior measure. The market consists of the benchmark instruments we use for the calibration. We denote with  $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)$  the vector of prices for the benchmark instruments and with  $\mathbf{G} = (G_1, \dots, G_K)$  the vector of their stochastic payoffs. With this in mind, we formulate the portfolio selection problem for the investor as

$$\max_{\boldsymbol{\phi}} \mathbb{E}^{\mathbb{P}}[u(\boldsymbol{\phi}^T \mathbf{G})] \quad \text{s.t.} \quad \boldsymbol{\phi}^T \boldsymbol{\pi} \leq W_0, \tag{15}$$

where  $\boldsymbol{\phi} = (\phi_1, \dots, \phi_K)$  is the portfolio choice. Since an unconstrained optimization problem generally allows more efficient computational methods, and the budget constraint can be assumed to hold with equality, we can rewrite (15) as

$$\max_{\boldsymbol{\phi}} \sum_{i=1}^N p_i u \left( \sum_{k=2}^K \phi_k \left( g_{ik} - g_{i1} \frac{\pi_k}{\pi_1} \right) + \frac{g_{i1} W_0}{\pi_1} \right), \tag{16}$$

where we substitute

$$\phi_1 = \frac{1}{\pi_1} \left( W_0 - \sum_{k=2}^K \phi_k \pi_k \right) \tag{17}$$

into the maximization problem above and  $g_{ik}$  denotes the payoff from option  $k$  when path  $i$  is realized. The first-order conditions are given by

$$\sum_{i=1}^N \left[ p_i \left( g_{ih} - g_{i1} \frac{\pi_h}{\pi_1} \right) u' \left( \sum_{k=2}^K \phi_k \left( g_{ik} - g_{i1} \frac{\pi_k}{\pi_1} \right) + \frac{g_{i1} W_0}{\pi_1} \right) \right] = 0, \tag{18}$$

for  $h = 1, \dots, K$ . The solution  $\boldsymbol{\phi}^* = (\phi_1^*, \dots, \phi_K^*)$  to this problem, which needs to be computed using numerical methods, is the optimal portfolio choice for the representative investor with initial wealth  $W_0$ .

The pricing kernel is now obtained by plugging  $(\phi^*)^\top G$  into  $u'(\cdot)$ , and the corresponding change of measure gives us the new risk neutral distribution we are after. More precisely, the new weights  $q_i$  for  $i = 1, \dots, N$  are given by

$$q_i = \frac{p_i u' \left( \sum_{k=2}^K \phi_k^* \left( g_{ik} - g_{i1} \frac{\pi_k}{\pi_1} \right) + \frac{g_{i1} W_0}{\pi_1} \right)}{\sum_{i=1}^N p_i g_{i1}} \pi_1. \tag{19}$$

The least squares setup is the following:

$$\max_{\phi} \sum_{i=1}^N p_i u \left( \sum_{k=2}^K \phi_k \left( g_{ik} - g_{i1} \frac{\pi_k}{\pi_1} \right) + \frac{g_{i1} W_0}{\pi_1} \right) - \frac{1}{2} \sum_{k=2}^K w_k \phi_k^2 - \frac{1}{2} \left( \frac{1}{\pi_1} \left( W_0 - \sum_{k=2}^K \phi_k \pi_k \right) \right)^2, \tag{20}$$

with the first-order conditions given by

$$\sum_{i=1}^N \left[ p_i \left( g_{ih} - g_{i1} \frac{\pi_h}{\pi_1} \right) u' \left( \sum_{k=2}^K \phi_k \left( g_{ik} - g_{i1} \frac{\pi_k}{\pi_1} \right) + \frac{g_{i1} W_0}{\pi_1} \right) \right] - w_h \phi_h + \frac{\pi_h}{\pi_1} \left( W_0 - \sum_{k=2}^K \phi_k \pi_k \right) = 0, \tag{21}$$

with  $h = 1, \dots, K$  and the weights  $q_i$  for  $i = 1, \dots, N$  given by (19) as before.

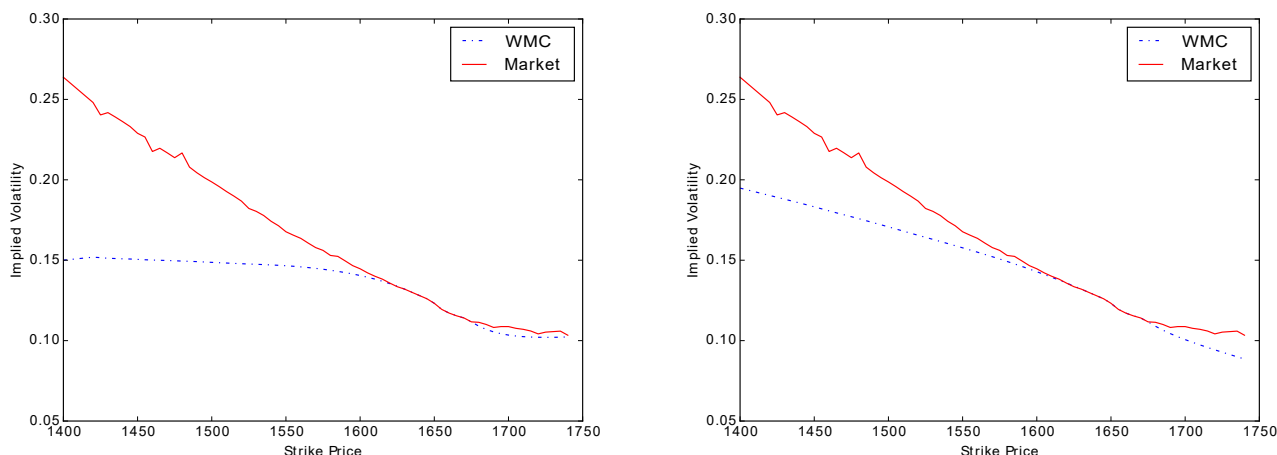
As previously mentioned, if we set  $u(x) = -e^{-x}$  and  $p_i = \frac{1}{N}$  for  $i = 1, \dots, N$ , the formulation is mathematically equivalent to the original weighted Monte Carlo method of Avellaneda et al. [13], where the measure of statistical distance is given by the Kullback–Leibler divergence and the prior measure is uniform. The initial wealth  $W_0$  in this case does not affect the solution, since  $u(x)$  is translation invariant. In the general case, a straightforward choice for  $W_0$  that conforms to the representative investor model is the “market wealth”, i.e., the value of the underlying assets.

#### 4. Calibration with Probability Distortion

##### 4.1. Introducing Risk Aversion and Probability Distortion

From the numerical results in Section 5, we learn that the accuracy of the original weighted Monte Carlo method is rather limited for out-of-sample options. For example, as can be seen in Figure 1, the implied volatility we obtain from the original weighted Monte Carlo method turns out to be much lower for the far-OTM options than what is implied by the market prices. From the consumption-based asset pricing perspective, this could in theory be explained by positing that the preferences represented by (13), and by extension of duality the relative entropy formulation in (3) are not risk averse enough. More specifically, (13) trivially includes a coefficient of risk aversion equal to one.

Hence, as a first attempt to improve the accuracy, we add an explicit coefficient of risk aversion to the formulation in (13). Deriving the corresponding generalized version of (6) using the Legendre transform then gives us a divergence measurement that contains a parameter which directly affects the tail thickness of the derived subjective probability measure represented by  $q$ , with higher values for the risk aversion coefficient translating to thicker tails. Preliminary numerical tests revealed, however, that the risk aversion coefficient by itself had barely a noticeable effect on the tail thickness for the range of values which still allowed a decent fit with the benchmark instruments.



**Figure 1.** Typical view of the volatility for SPX options implied by market prices and by the WMC calibrated GBM model (left) or Heston model (right).

From a decision-theoretic perspective, we can, however, accommodate the thick tails implied by the market prices by positing that the market overweights the probability of large market movements with respect to the initial model. To implement this idea, we apply a probability distortion that is partly similar to that introduced by Tversky and Kahneman [19] in their work on CPT. This generalization turns out to give vastly better empirical performance than the original method as well as its purely utility function-based risk aversion modifications in our tests.

The general CPT specification is given by

$$U(X) = \int_0^{+\infty} W^+(\mathbb{P}(u^+(X) > x))dx - \int_0^{+\infty} W^-(\mathbb{P}(u^-(X) > x))dx, \tag{22}$$

where  $W^+, W^-, u^+$  and  $u^-$  are the probability distortion functions and utility functions over deterministic outcomes for the gain and loss domains, respectively, and  $\mathbb{P}$  is a probability measure. The distortion function we chose to implement for the numerical tests is the one introduced by Prelec [31], which gave considerably better results than the original distortion function in [19], and is given by

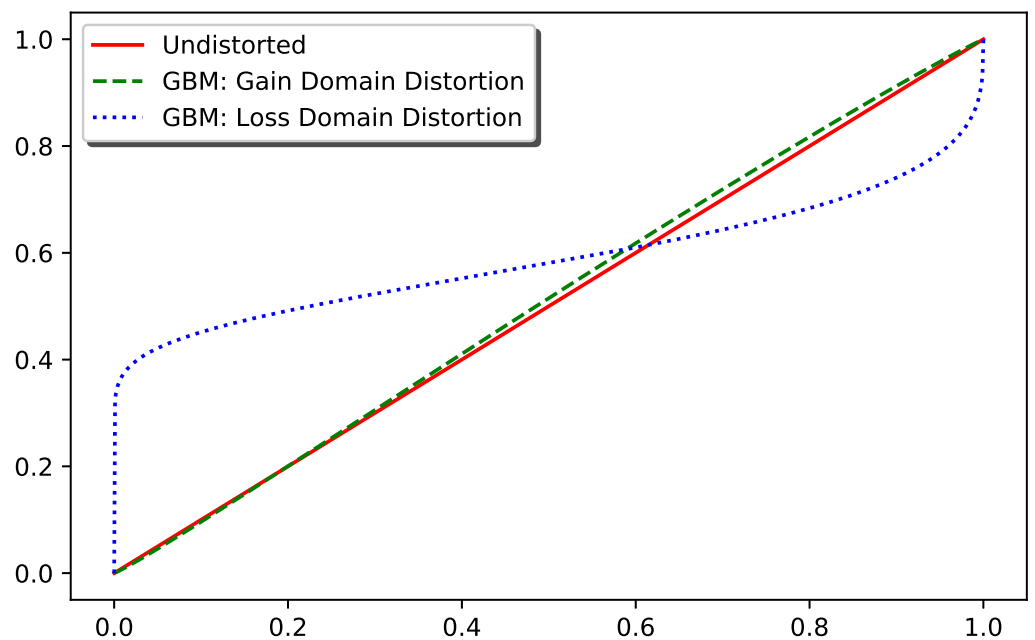
$$v^+(P) = \exp\{-\gamma^+(-\ln(P))^{\delta^+}\} \tag{23}$$

for the gains domain, and

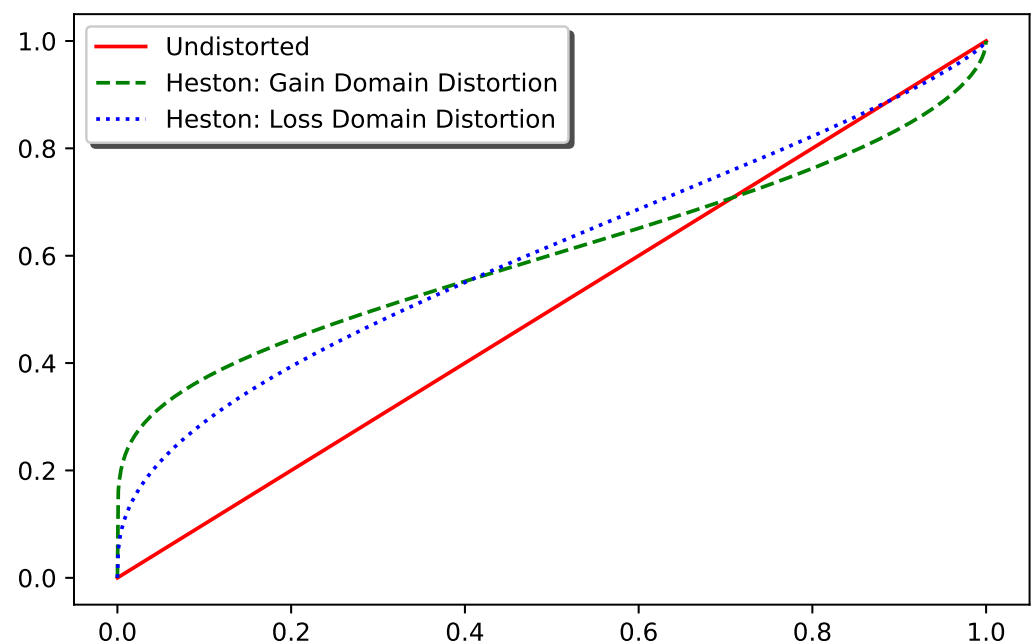
$$v^-(P) = \exp\{-\gamma^-(-\ln(P))^{\delta^-}\} \tag{24}$$

for the loss domain. Here,  $\delta^+$  and  $\delta^-$  correspond to the curvature of the distortion, while  $\gamma^+$  and  $\gamma^-$  correspond to the elevation of the distortion, with a value of 1 for all parameters corresponding to a non-distorted probability measure (see [31] for a comprehensive discussion on the interpretation of these parameters). Figures 2 and 3 illustrate the distortion functions used in Section 5.





**Figure 2.** Distortion functions for the GBM model with mean probability distortion values  $\gamma = 1.051$  and  $\delta = 0.97$  for the gain domain and  $\gamma = 0.316$  and  $\delta = 0.611$  for the loss domain.



**Figure 3.** Distortion functions for the Heston model with mean probability distortion values  $\gamma = 0.555$  and  $\delta = 0.623$  for the gain domain and  $\gamma = 0.791$  and  $\delta = 0.640$  the loss domain.

We can adapt the formulation in (22) to a discrete state space in the following way. Let  $\Omega = \{x_{-m}, \dots, x_0, \dots, x_n\}$  be the set of possible outcomes, such that  $x_{-m} \leq x_{-m+1} \leq \dots \leq x_{n-1} \leq x_n$  with  $x_0 = 0$  by convention, and let  $p_i$  be the probability of outcome  $x_i$  for  $i = -m, \dots, n$ . Furthermore, let  $u^- : \Omega^- \rightarrow \mathbb{R}$  be a strictly increasing convex function, with  $\Omega^- = \{x_{-m}, \dots, x_{-1}\}$ , and let  $u^+ : \Omega^+ \rightarrow \mathbb{R}$  be a strictly increasing concave function, with  $\Omega^+ = \{x_0, \dots, x_n\}$ . Furthermore, let  $v^- : [0, 1] \rightarrow [0, 1]$  and  $v^+ : [0, 1] \rightarrow [0, 1]$  be two strictly increasing functions such that  $v^-(0) = v^+(0) = 0$  and  $v^-(1) = v^+(1) = 1$ . If we

denote by  $X = (x_{-m}, p_{-m}; \dots; x_{-1}, p_{-1}; x_0, p_0; x_1, p_1; \dots; x_n, p_n)$  the prospect  $X$ , then the cumulative prospect value of  $X$  is given by

$$U(X) = \sum_{i=-m}^n v_i u(x_i), \tag{25}$$

where

$$u(x_i) = \begin{cases} u^+(x_i), & \text{if } 0 \leq i \leq n, \\ u^-(x_i), & \text{if } -m \leq i < 0, \end{cases} \tag{26}$$

and

$$v_i = \begin{cases} v^+(p_i + \dots + p_n) - v^+(p_{i+1} + \dots + p_n), & \text{if } 0 \leq i \leq n, \\ v^-(p_{-m} + \dots + p_i) - v^-(p_{-m} + \dots + p_{i-1}), & \text{if } -m \leq i < 0. \end{cases} \tag{27}$$

#### 4.2. The Weighted Monte Carlo Method with Probability Distortion

We now turn to the main contribution of this paper, which is a method that combines probability distortion in style of CPT preferences with the weighted Monte Carlo method. This allows us to adapt the no-arbitrage model we want to calibrate to the thick tails implied by far-OTM options while maintaining an exact fit with the benchmark instruments we use for the calibration.

When the benchmark instruments all have the same maturity, we proceed exactly as in Section 2 where we solve (6) to obtain the new weights  $q$  with the exception that we now use a prior measure  $p$  which incorporates probability distortion that re-weights the tail events of the initial model. This is in contrast to the original weighted Monte Carlo method where the prior measure is simply taken to be uniform.

We define the tail events to be large changes in the price of the underlying asset. That means we need to sort the realized paths according to their value at the time the benchmark options expire. We then enumerate them in an ascending order as  $-m, -m + 1, \dots, n - 1, n_1$ , where path  $i = 0$  is the closest to realizing no change in value for the underlying among the sampled paths, and compute the prior measure as

$$p_i = \begin{cases} v^+(\frac{n-i}{N}) - v^+(\frac{n-i-1}{N}), & \text{if } 0 \leq i \leq n, \\ v^-(\frac{m-i}{N}) - v^-(\frac{m-i-1}{N}), & \text{if } -m \leq i < 0, \end{cases} \tag{28}$$

where  $v^+(\cdot)$  and  $v^-(\cdot)$  are the distortion functions for the gain and loss domains, respectively.

As mentioned in the Introduction, a portfolio choice problem with CPT preferences generally leads to a non-additive expectation operator, which is not the case in the formulation above. To understand why we avoid a non-additive expectation we recall that the Choquet integral is additive for comonotonic random variables (see, e.g., [32] for an in-depth discussion of these concepts). For a portfolio choice problem where the available instruments consist of an underlying asset and options on that asset the prospect outcomes can most generally be taken to be the value of the portfolio of the investor, which generally is not comonotonic with the underlying asset if, for example, the portfolio includes short positions on that asset. However, in our formulation, the probability distortion does not depend on  $\lambda$  (or  $\phi$  in the utility maximization case), so the expectation is trivially contained within a single comonotonic class.

If the algorithm is implemented in such a way that the portfolio problem is solved for only one option maturity, the ordering is straightforward, since the value of each realized path relative to the rest is unambiguous at maturity. If more than one benchmark maturity is present, the ordering of entire paths is no longer unambiguous, since the simulated paths can cross each other between maturities. The method proposed here tackles this issue by splitting the set of benchmark instruments into single-maturity subsets, and solving the portfolio problem for each subset separately. In other words, if our benchmark instruments consist of options with  $M$  different maturities, we split the options into  $M$  groups and solve the equivalent of  $M$  single maturity entropy minimization problems. This results in

each path being assigned a vector of weights. Realized values along the simulated paths at times that do not coincide with the maturities present in the set of benchmark options are then assigned weights based on linear interpolation with respect to their relative position. This is in contrast with the original weighted Monte Carlo method which assigns a single weight to entire paths.

The full specification of the method we propose is as follows. Let  $J_b$  and  $J_s$  be the two sets of indices such that  $\Gamma_b = \{t_j | j \in J_b\}$  is the set of all points in time between 0 and  $T$  that coincide with the maturity of a benchmark option, and  $\Gamma_s = \{t_j | j \in J_s\}$  is the set of all points in time between 0 and  $T$  that do not coincide with a benchmark maturity but for which we would like to know the state price density implied by the benchmark options. Furthermore, given  $j \in J_b$ , let  $S^j = \{x_{-m}^j, \dots, x_{-1}^j, x_0^j, x_1^j, \dots, x_n^j\}$  denote the ordered set of realized points in our sample space at time  $t_j$ , such that  $x_i^j < x_{i'}^j$  for  $i < i'$ .

We start by computing the prior weights  $p^j$  using (28). Next, we compute the new weights  $q^j$  by first solving either (6) for an exact fit or (8) for an approximate fit and then plugging the solution  $\lambda^j$  into (5). Once the weights have been computed for each maturity  $t_j \in \Gamma_b$ , the set of weights for each  $t_{j'} \in \Gamma_s$  is calculated as follows. Assume that we have  $j, j'' \in J_b$  and  $j' \in J_s$  such that  $t_j = \sup\{\tau \in \Gamma_b | \tau < t_{j'}\}$  and  $t_{j''} = \inf\{\tau \in \Gamma_s | \tau > t_{j'}\}$ . Then, we calculate the sort index arrays  $I^j, I^{j'}$ , and  $I^{j''}$  (i.e., the index arrays for the sorted versions of  $S^j, S^{j'}$ , and  $S^{j''}$ ). Denote the weight arrays for  $S^j$  and  $S^{j''}$  by  $q^j$  and  $q^{j''}$ , and for the sake of visual clarity let us introduce the notation  $q^j(i) \equiv q_i^j$ . If  $I^j(i) = i'$ , we compute

$$q^{j'}(i') = \frac{t_{j''} - t_j}{t_{j''} - t_j} q^j(I^j(i')^{-1}) + \frac{t_{j'} - t_j}{t_{j''} - t_j} q^{j''}(I^{j''}(i')^{-1}), \tag{29}$$

for  $i = 1, \dots, N$ . Finally, if either  $t_j$  or  $t_{j''}$  does not exist, i.e., if the maturity of  $t_{j'}$  falls outside the range of maturities of the benchmark instruments, then we simply choose whatever benchmark maturity is closest and put the interpolation weight of that maturity to one.

To summarize, the calibration algorithm proposed above works in its general form as follows:

1. A set of  $N$  paths is generated from the initial model, and the payoff matrix in (2) is computed.
2. The paths are indexed according to their sort order at each of the  $M$  benchmark maturities, and the set of prior weights  $p^j$  for  $j = 1, \dots, M$  is computed.
3. With these payoffs and prior weights in hand, we solve separately for each  $t_j$  the problem given by (6) (or (11) for an approximate fit) using the prior weights  $p^j$  and the restriction of the payoff matrix to the set of benchmark instruments which expire at  $t_j$ . We then plug the solution  $\lambda^j$  into (5) to obtain  $q^j$  which become the weights attached to the points on the sample paths at time  $t_j$ .
4. For a maturity  $t_{j'}$  of interest which does not coincide with the maturities present in the set of benchmark instruments, we use the formula given by (29) to interpolate between the two weights vectors  $q^j$  and  $q^{j''}$  obtained for the two subsets of benchmark instruments whose maturities  $t_j$  and  $t_{j''}$  most narrowly sandwich  $t_{j'}$ . In the event that the maturity is either longer or shorter than anything that is available in the set of benchmark instruments, we simply attach to it the weights vector computed for the set of benchmark instruments with the maturity closest to  $t_{j'}$ .

We see that, if we set the probability distortion function in (28) equal to the identity function and drop the distinction between benchmark instruments based on their maturity, then we retrieve the original weighted Monte Carlo method. For the sake of conciseness, we hereafter refer to the procedure given above as the generalized weighted Monte Carlo (GWMC) method and to the special case where no partition is performed on the set of benchmark instruments and where the prior measure is uniform as the original weighted Monte Carlo (OWMC) method.

### 4.3. Path Dependent Option Pricing with GWMC-Calibrated Paths

In the most basic setup of the weighted Monte Carlo method, where a single weight is assigned to each path, the probabilistic interpretation is clear. The weight  $q_i$  represents the subjective probability that path  $i$ , in the state space that consists of the sample paths generated by the initial model, is realized. However, the GWMC method gives us several weights per path. The question then becomes how we interpret the multiple weights per path, and how they allow us to compute the expected payoff.

We start by realizing that, when there are  $M$  maturities present, we can view the GWMC model as a sequence of  $M$  one period models, where the unconditional subjective probability distribution  $q^m = \{q_1^m, \dots, q_N^m\}$  for maturity  $\tau_m$  with  $m \in \{1, \dots, M\}$  is what we would get if we only solved the utility maximization problem, or the equivalent entropy minimization problem, using only the benchmark options with maturity  $\tau_m$ .

What we really need, however, is the distribution  $q = \{q_1, \dots, q_N\}$  that describes the probability of observing the entire path  $i$  for  $i = 1, \dots, N$ . How we get  $q$  from  $\{q^1, \dots, q^M\}$  is revealed by the following theorem.

**Theorem 1.** Let  $s_i = \{s_i^1, \dots, s_i^M\} \in \Omega$  denote the  $i$ th realization of  $S = \{S^j\}_{j=1}^M$  in our sample of  $N$  paths. Furthermore, assume that for each  $h, i = 1, \dots, N$  and  $j, m = 1, \dots, M$  we have that  $s_i^j = s_h^m$  if and only if  $i = h$  and  $j = m$ , i.e., no sample path  $i$  in  $S$  contains any elements also contained by another sample path  $h$  of  $S$ . Then, the probability of path  $i$  being realized under the subjective measure  $\mathbb{Q}$  is given by

$$\mathbb{Q}(\{S = s_i\}) = \frac{1}{M} \sum_{j=1}^M q_i^j, \tag{30}$$

where  $q_i^j$  is the unconditional probability under  $\mathbb{Q}$  of observing  $s_i^j$ .

**Proof.** By the law of total probability, we have that

$$\mathbb{Q}(\{S = s_i\}) = \sum_{j=1}^M \mathbb{Q}(\{S = s_i | s_i^j\}) \mathbb{Q}(\{s_i^j\}),$$

where  $\mathbb{Q}(\{s_i^j\})$  is the unconditional probability that  $s_i^j$  is realized at all. We have that  $s_i^j$  completely identifies the path  $s_i$ , by our assumption of sample uniqueness. Therefore,  $\mathbb{Q}(\{S = s_i | s_i^j\}) = 1$ , and we can write

$$\mathbb{Q}(\{S = s_i\}) = \sum_{j=1}^M \mathbb{Q}(\{s_i^j\}) = \sum_{j=1}^M \frac{1}{M} q_i^j.$$

Note that  $q_i^j$  is the subjective probability of observing  $s_i^j$  at time  $t_j$ . However, the unconditional subjective probability of observing  $s_i^j$  at all is given by

$$\mathbb{Q}(\{s_i^j\}) = \frac{q_i^j}{M} \tag{31}$$

since  $s_i^j$  can only appear at time  $t_j$ , and not in any of the other  $M - 1$  maturities, by our assumption of unique realizations of  $S$ .  $\square$

In other words, the subjective probability of observing a given sample path when the weights are determined by the GWMC is simply the average of the weights along that path. Note that our assumption of uniqueness among the values generated by the simulation of the initial model is a simplifying one, but we would expect it to hold for any decent random number generator. Assuming uniqueness is well justified in any case. If

we need to resort to simulation methods to begin with, it is quite safe to assume that the distribution of the underlying asset has continuous support, which means that repeated values in a finite sample generated in an unbiased way from the distribution should be a zero-measure event. In other words, the probability that we draw the same value more than once from the distribution of the underlying should be zero, almost surely. When we talk about a “decent” random number generator in this context, we simply mean that it imitates the underlying (continuous) distribution well enough that repeated values are not encountered.

We conclude this section by illustrating the application of the GWMC to path dependent options by the way of an example. Consider the Monte Carlo formula for the price of an arithmetic Asian option with  $M$  monitoring points. The standard Monte Carlo price for this option is given by

$$\pi_A = \frac{1}{N} \sum_{i=1}^N \left| \frac{1}{M} \sum_{t=1}^M s_i^j - X \right|^+, \tag{32}$$

with  $X$  being the strike price. Now, let  $q_i^j$  denote the weight of path  $i$  at time  $t_j$ . If we write

$$\bar{q}_i = \frac{1}{M} \sum_{j=1}^M q_i^j, \tag{33}$$

the price  $\pi_A$  is given by

$$\pi_A = \sum_{i=1}^N \bar{q}_i \left| \frac{1}{M} \sum_{j=1}^M s_i^j - X \right|^+. \tag{34}$$

If all the weights  $q_i^j$  are identical for a given  $i$ , this expression simplifies to the original weighted Monte Carlo formulation. If we further set  $q_i^j = \frac{1}{N}$  for every  $i$  and every  $j$ , we get the usual Monte Carlo pricing formula for the arithmetic Asian option.

### 5. Implementation Details and Numerical Results

Our numerical experiments consisted of two parts:

1. A cross-sectional run where we tested the GWMC method for single maturities only
2. An intertemporal run where we tested the GWMC method for multiple maturities simultaneously

In Section 5.1, we describe the initial models we used and the pre-calibration we employed to estimate their parameters, as well as the probability distortion function we chose to implement. In Sections 5.2 and 5.3, we give the numerical results from the cross-sectional and intertemporal tests, respectively.

Our numerical tests were performed on SPX options priced during the period from 1 January 2013 to 31 December 2013. This dataset contains a total of 765,952 contracts. We calculated the risk free rate by linearly interpolating yields from US Treasury bills data available on the US Federal Reserve website. The dividend payments on the S&P 500 index were approximated by a continuous dividend yield. In addition to the options, we included the underlying asset itself as well as a risk-free asset in the set of benchmark instruments.

Throughout our empirical tests, we calculated two types of error measurements: the mean relative price error (MRPE) and the mean average price error (MAPE). More precisely,

$$MRPE = \frac{1}{K} \sum_{k=1}^K \frac{|\pi_k^{Model} - \pi_k^{Market}|}{\pi_k^{Market}}, \tag{35}$$

and

$$MAPE = \frac{1}{K} \sum_{k=1}^K |\pi_k^{Model} - \pi_k^{Market}|, \tag{36}$$

where  $\pi_k^{Model}$  is the model-predicted price of benchmark instrument  $k$  and  $\pi_k^{Market}$  is the price of that instrument observed in the market.

The data exhibit a significant number of put–call parity violations, which can likely be attributed to the fact that we used end-of-day prices, which leads to a degree of asynchronicity. For this reason, all of the numerical tests were done on out-of-money puts and out-of-money calls. As the number of benchmark instruments increases, the entropy minimization/utility maximization part of the calibration procedure can become a challenge in itself, particularly when an exact solution is sought. We kept the number of benchmark instruments small for this reason and used the least squares approach in the intertemporal tests. In addition, we performed each calibration separately for the puts and the calls to further reduce calibration failures.

For the pre-calibration (i.e., the estimation of the parameters of the Black–Scholes and Heston models), we used Matlab’s `lsqnonlin` routine, and, for the calibration runs, we used the BFGS routine in Python’s Scipy library (see [33] for a discussion on the computational considerations on the weighted Monte Carlo method).

### 5.1. Initial Models, Pre-Calibration and Path Generation

We used two no-arbitrage models as a prior for our calibration procedure: the Black–Scholes [1] model based on geometric Brownian motion (GBM) and the Heston [3] stochastic volatility model. Sampling with these models was done using the Euler discretization scheme. More specifically, for the GBM model, we generated the paths using the discretized dynamics

$$S_{t+\Delta t} = S_t e^{(r-d-\frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}Z}, \tag{37}$$

where  $r$  is the risk-free rate,  $d$  is the continuous dividend yield,  $\sigma$  is the volatility, and  $Z$  are standard Gaussian innovations, i.e.,  $Z \sim N(0, 1)$ . For the Heston model, we generated the paths using

$$\begin{aligned} S_{t+\Delta t} &= S_t e^{(r-d-\frac{1}{2}\sigma^2)\Delta t + \sqrt{v_t\Delta t}Z_S}, \\ v_{t+\Delta t} &= v_t + \kappa(\theta - v_t)\Delta t + \eta\sqrt{v_t\Delta t}Z_v, \end{aligned} \tag{38}$$

where  $S_t$  is the price process of the underlying asset,  $v_t$  is the variance process,  $r$  and  $\delta$  are the risk free rate and dividend yield as before,  $\kappa$  is the rate of mean reversion,  $\theta$  is the long run volatility, and  $\eta$  is the volatility-of-volatility. Here, the innovations  $Z_v \sim N(0, 1)$  and  $Z_S \sim N(0, 1)$  are correlated with correlation coefficient  $\rho$ .

The calibration runs for the out-of-sample performance tests were all done using  $N = 40,000$  simulated paths with antithetic variance reduction. For the Heston model, we used a full truncation scheme for the variance process to prevent it from becoming negative. Here, full truncation means the variance process is given by  $\max\{V_t, 0\}$  at every sample time  $t$ . Each path simulated for the Heston model contained 100 points which were distributed equally between each benchmark maturity present along the path.

The pre-calibration of these models was done using a nonlinear least squares approach. That is, the parameter values  $\chi^*$  for the respective models were found by computing

$$\chi^* = \underset{\chi}{\operatorname{argmin}} \left\{ \sum_{k=1}^{K-2} w_k \left( \pi_k^{Model} - \pi_k^{Market} \right)^2 \right\}, \tag{39}$$

where the index only reaches  $K - 2$  since we performed the pre-calibration using only the option data, leaving out the underlying and risk-free assets added for the subsequent weighted Monte Carlo calibration tests.

The weights  $w_k, k = 1, \dots, K - 2$  were calculated as the inverse of the bid–ask spread for the corresponding option, i.e.,

$$w_k = \frac{1}{\pi_k^{Market,ask} - \pi_k^{Market,bid}}. \tag{40}$$

For the geometric Brownian model, the decision variable is  $\chi = \sigma$ , whereas for the Heston model we have  $\chi = (v_0, \theta, \eta, \kappa, \rho)$ . In terms of European option pricing, both the geometric Brownian motion model and the Heston model have closed form solutions, although in the latter case we make use of the characteristic function, which contains an integral which must be evaluated numerically.

A pre-calibration was done for each open market day over the period of 1 January 2013 to 31 December 2013. For each such day, the set of options used in the pre-calibration procedure consisted of the 14 most traded out-of-money puts, together with the 14 most traded out-of-money calls for each maturity. For maturities where fewer than 14 puts (calls) were traded that day, we simply included all put (call) options with nonzero trading volume. The summary of the parameter estimates obtained from this pre-calibration procedure is given in Table 1.

**Table 1.** Pre-calibration summary for initial models.

	$\sigma$	$v_0$	$\theta$	$\eta$	$\rho$	$\kappa$
mean	0.1313	0.0166	0.0422	0.3981	−0.9337	1.9219
std	0.0063	0.0030	0.0039	0.0313	0.0207	0.4448
max	0.1435	0.0232	0.0477	0.4655	−0.8979	2.8056
min	0.1237	0.0127	0.0341	0.3514	−0.9715	1.3233

## 5.2. Cross-Sectional Calibration Results

Our goal for this part of the numerical tests was to include as much of the strike range of the out-of-money puts and calls for each maturity as possible, to see the full effect of the over-and-underweighting of probabilities of extreme events on the pricing measure. However, at the extreme ends of option moneyness, the data become noticeably less reliable, with instances of duplicate prices for options with different strikes, bid prices of zero, and so on. For these reasons, we removed the following:

- Any option with bid price zero
- Any option with price lower than 0.5
- Any set of options with different strikes but the same price, same option type, and same maturity quoted on the same day
- Any option of which the open interest count fell below 2000 for short and medium maturities and 1000 for long maturities
- Any set of options of the same maturity quoted on a given trading day with fewer than 10 options satisfying the above criteria

Here, short maturities are defined to be between 1 and 90 days, medium maturities between 91 and 250 days, and long maturities anything longer than 250 days. We used open interest as our measure of liquidity, instead of trading volume, since trading volume gave a much thinner support for options at the far ends of the moneyness spectrum, as well as for options with long maturities.

After the aforementioned contracts had been removed, we were left with a set of 24,263 contracts, consisting solely of out-of-money puts and calls. For this set of contracts, we performed an out-of-sample test for three different calibration methods for two different initial models:

1. The unweighted geometric Brownian motion model
2. The unweighted Heston model
3. The geometric Brownian motion model calibrated with the OWMC
4. The Heston model calibrated with the OWMC
5. The geometric Brownian motion model calibrated with the GWMC
6. The Heston model calibrated with the GWMC

For each model–method pair, we calculated the out-of-sample performance for each maturity and option type (i.e., puts or calls) separately, as well as their aggregate perfor-

mance over these categories. These results are given in Tables 2 and 3, while Table 4 gives the aggregate result for each model–method pair. The first entry in each number pair in these tables is the mean relative pricing error, and the second entry is the mean absolute pricing error, as explained in Section 5.1. For each option type (i.e., put or call), we used as benchmark instruments the five options with strikes closest to the forward price of the underlying, so a total of 10 benchmark options per maturity. The remainder of the options in the dataset, obtained from the data cleaning procedure described above, were used as out-of-sample instruments.

**Table 2.** Cross-sectional results for the unweighted version, the OWMC method, and the GWMC method using geometric Brownian motion as an initial model, for OTM puts ( $K < F$ ) and OTM calls ( $K > F$ ). The first and second numbers in each column are the MRPE and MAPE, respectively.

Method	Moneyness	Short	Medium	Long	All
Unweighted	K <F	0.624, 3.216	0.818, 9.492	0.853, 24.520	0.748, 11.286
	K >F	0.855, 3.515	0.530, 3.429	0.345, 3.630	0.690, 3.513
	All	0.684, 3.294	0.768, 8.442	0.789, 21.901	0.736, 9.727
OWMC	K <F	0.439, 1.071	0.506, 2.444	0.444, 5.625	0.457, 3.042
	K >F	0.121, 0.275	0.089, 0.295	0.076, 0.438	0.101, 0.331
	All	0.355, 0.860	0.435, 2.078	0.369, 4.567	0.378, 2.439
GWMC	K <F	0.186, 0.331	0.224, 0.567	0.208, 1.343	0.203, 0.755
	K >F	0.072, 0.155	0.079, 0.264	0.068, 0.391	0.072, 0.249
	All	0.155, 0.284	0.198, 0.514	0.180, 1.154	0.174, 0.643

**Table 3.** Cross-sectional results for the unweighted version, the OWMC method, and the GWMC method using the Heston model as an initial model, for OTM puts ( $K < F$ ) and OTM calls ( $K > F$ ). The first and second numbers in each column are the MRPE and MAPE, respectively.

Method	Moneyness	Short	Medium	Long	All
Unweighted	K <F	0.234, 0.798	0.362, 2.861	0.391, 8.284	0.317, 3.592
	K >F	0.292, 1.569	0.314, 2.356	0.297, 4.283	0.298, 2.215
	All	0.249, 1.001	0.353, 2.773	0.379, 7.782	0.313, 3.316
OWMC	K <F	0.271, 0.544	0.222, 0.780	0.152, 1.138	0.216, 0.815
	K >F	0.216, 0.397	0.270, 0.813	0.239, 1.210	0.232, 0.730
	All	0.256, 0.505	0.230, 0.787	0.170, 1.153	0.220, 0.796
GWMC	K <F	0.084, 0.206	0.049, 0.195	0.067, 0.593	0.067, 0.322
	K >F	0.075, 0.147	0.044, 0.120	0.042, 0.161	0.060, 0.142
	All	0.081, 0.190	0.048, 0.182	0.063, 0.536	0.066, 0.286

**Table 4.** Aggregate results for the out-of-sample cross-sectional performance for the two different initial models along with the two different weighted calibration methods, as well as the unweighted versions for comparison. The first and second number in each column are the average MRPE and MAPE, respectively, for the entire dataset.

Method	GBM	Heston
Unweighted	0.736, 9.727	0.313, 3.316
OWMC	0.378, 2.439	0.220, 0.796
GWMC	0.174, 0.643	0.066, 0.286

For the calibration of a given trading day, we used the distortion parameter values,  $\xi = \{\gamma^+, \delta^+, \gamma^-, \delta^-\}$ , which gave the best out-of-sample performance during the calibration of the previous trading day. As can be seen from comparing the values in Table 5, we expect these values to be different for different initial models. More specifically, before calibrating



the model for a given trading day, we solve the following for the trading day immediately before it:

$$\zeta^* = \underset{\zeta}{\operatorname{argmin}} \left\{ \frac{1}{K} \sum_{k=1}^K \frac{|\pi_k^{\text{Model}}(\zeta) - \pi_k^{\text{Market}}|}{\pi_k^{\text{Market}}} \right\}. \tag{41}$$

This optimization problem was solved using BFGS. While each function evaluation in this minimax optimization problem is expensive (we are solving the portfolio choice problem with each function call), the computational times are drastically reduced by reusing the previous optimal parameter estimate as a starting point, since the optimal values turn out to change very little between trading days in our data. In addition, since there is no interaction between the gain domain parameters and the loss domain parameters, they can be calibrated separately, which effectively halves the time complexity of the calibration problem, which is given by  $\mathcal{O}(n^2)$  for the BFGS routine, where  $n$  is the number of decision variables. On average, a single cross-sectional run took 0.78 s for the OWMC method and 1.3 s for the GWMC. The calibration of the parameters in  $\zeta$  took roughly 11 min on average.

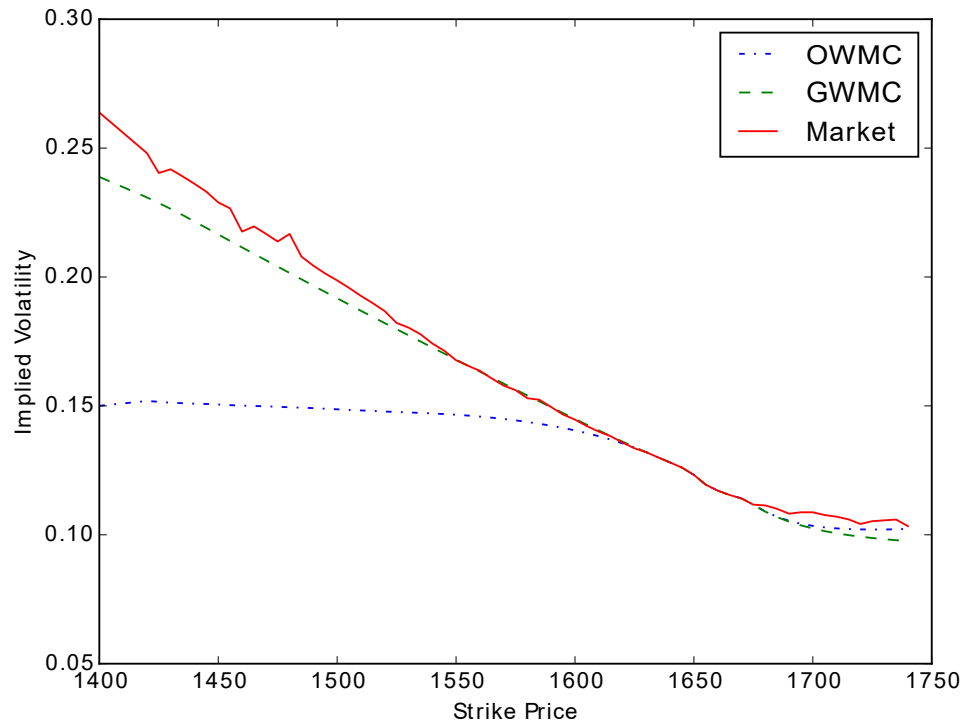
**Table 5.** The averages of the probability distortion parameter values for different maturity clusters and models, along with their standard deviations in brackets.

Model	Horizon	$\gamma^+$	$\delta^+$	$\gamma^-$	$\delta^-$
GBM	Short	1.137 (0.124)	0.921 (0.132)	0.431 (0.129)	0.577 (0.178)
	Medium	1.051 (0.138)	0.975 (0.166)	0.316 (0.111)	0.611 (0.142)
	Long	1.100 (0.083)	1.022 (0.090)	0.332 (0.105)	0.580 (0.107)
Heston	Short	0.614 (0.149)	0.588 (0.168)	0.630 (0.153)	0.547 (0.177)
	Medium	0.555 (0.157)	0.623 (0.182)	0.791 (0.139)	0.640 (0.210)
	Long	0.451 (0.112)	0.601 (0.124)	0.788 (0.116)	0.754 (0.118)

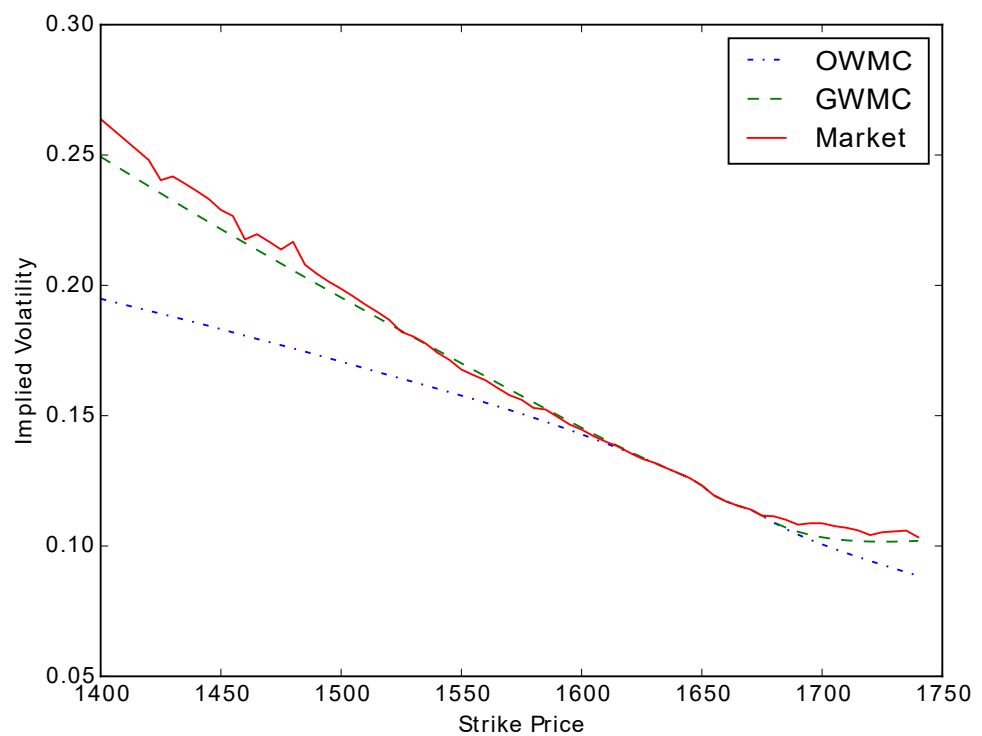
Figures 4 and 5 give an idea of how the GWMC improves upon the original weighted Monte Carlo method in terms of the empirical fit. They show a cross-section of the implied volatility calculated from market prices of options traded on 21 May 2013 with maturity of 30 days, plotted together with the OWMC and GWMC implied volatilities for those same options. The benchmark options for the OWMC and GWMC consisted of five close to the money puts and five close to the money calls. The figures demonstrate a typical difference between the OWMC and the GWMC as pricing models; the former tends to underprice far-out-of-the-money options to a much greater degree than the latter, with the exception of call options for geometric Brownian motion as the initial model.

When comparing the performance of a given model across the three maturity categories, it should be kept in mind that the presence of options of extreme moneyness tended to be significantly more prevalent in the medium range of our data than in the other two maturity groups, in addition to the fact that the liquidity requirements were not the same across the maturity groups. With that said, the Heston model outperforms the GBM model overall as expected in the unweighted case. Interestingly, though, the OWMC for GBM appears to achieve very good performance for the call options, which highlights the fact that a poor empirical fit of the initial model does not necessitate a poor fit once the paths have been reweighted. Overall, however, the GWMC method produces significant improvements, both for the geometric Brownian motion and the Heston model compared to the OWMC. It should also be noted here that one reason the weighted GBM models

underperform, particularly for the put options, is that the Monte Carlo simulation often did not produce any paths that reached the extreme levels necessary for the most far-out options to become exercised, while the Heston model did. The distortion function does not take effect for zero-measure events.



**Figure 4.** The volatility for SPX options with maturity 30 days implied by market prices, the OWMC calibrated GBM model, and the GWMC calibrated GBM model.



**Figure 5.** The volatility for SPX options with maturity 30 days implied by market prices, the OWMC calibrated Heston model, and the GWMC calibrated Heston model.

### 5.3. Intertemporal Calibration Results

The numerical tests presented in this section consisted of calibrating the initial models of Section 5.1 to benchmark options spanning more than one maturity. For each benchmark option maturity present during a given trading day, we chose the 10 out-of-money put options and the 10 out-of-money call options with the highest trading volume. Out of these, we used the five out-of-money put options and the five out-of-money call options that had strikes closest to the forward value as benchmark options. The remainder of the options served as out-of-sample options. We used the same probability distortion values as are given in Table 5 in Section 5.2.

For the purpose of demonstration, let  $U_d$ , with  $d = 1, \dots, D$  denote the set of all options for trading day  $d$ , where  $D$  is the final day (i.e., 31 December 2013), in our sample with nonzero trading volume. Furthermore, for each  $U_d$ , let  $M_{d,j} \subseteq U_d$  denote the set of the 10 puts and 10 calls with the highest trading volume with maturity  $m_j$ , where the maturities are indexed here by  $j = 1, \dots, J$  in such a way that if  $j' < j''$  then  $m_{j'} < m_{j''}$ . In addition, let  $K_{d,j} \subseteq M_{d,j}$  be the set of five puts and five calls in  $M_{d,j}$  that are closest to being at-the-money. Lastly, define the Cartesian product  $H_d := \otimes_{j=1}^J M_{d,j}$  and let  $\text{Dim}(H_d) := J$  be the number of different maturities present in  $H_d$ . The intertemporal test can then be described in the following way. For a given trading day  $d$ :

1. If  $\text{Dim}(H_d) \leq 3$ , the trading day was dropped from the sample.
2. If  $\text{Dim}(H_d) = 4$ , the in-sample instruments consisted of  $K_{d,1}$  and  $K_{d,3}$  and the out-of-sample instruments consisted of  $M_{d,1} \setminus K_{d,1}$ ,  $M_{d,2}$ ,  $M_{d,3} \setminus K_{d,3}$ , and  $M_{d,4}$ .
3. If  $\text{Dim}(H_d) = 5$ , the in-sample instruments consisted of  $K_{d,2}$  and  $K_{d,4}$  and the out-of-sample instruments consisted of  $M_{d,1}$ ,  $M_{d,2} \setminus K_{d,2}$ ,  $M_{d,3}$ ,  $M_{d,4} \setminus K_{d,4}$ , and  $M_{d,5}$ .
4. If  $\text{Dim}(H_d) \geq 6$ , the in-sample instruments consisted of  $K_{d,2+\iota}$  and  $K_{d,5+\iota}$  and the out-of-sample instruments consisted of  $M_{d,1+\iota}$ ,  $M_{d,3+\iota} \setminus K_{d,3+\iota}$ ,  $M_{d,3+\iota}$ ,  $M_{d,4+\iota}$ ,  $M_{d,5+\iota} \setminus K_{d,5+\iota}$ , and  $M_{d,6+\iota}$  with  $\iota = 0, \dots, \text{Dim}(H_d) - 6$ .

To elaborate on Part 4, for  $H_d$  with  $\text{Dim}(H_d) = J > 6$ , we begin by setting  $\iota = 0$  and calibrate. Once the calibration is done and we have calculated the out-of-sample prices, if  $\text{Dim}(H_d) > 6$ , we set  $\iota = 1$ , which shifts the maturity index by one to the right to perform the calibration and out-of-sample calculation again, and so on, until  $M_{d,6+\iota} = M_{d,J}$ . This test scheme resulted in the calculation of 135,792 option prices. Most trading days fell under the fourth case, meaning most options were priced several times using different benchmarks, and the number reflects every such calculation. On average, a single intertemporal run took 0.38 s for the OWMC method and 1.1 s for the GWMC.

The design of the intertemporal test was made with the aim of avoiding biases in the selection of the benchmark options by including an approximately even mix of “outside” and “in between” options, and by rolling over every maturity as in Step 4 for each trading day. As the results in Tables 6 and 7 show, the Heston model gives better results overall than the geometric Brownian motion, and the GWMC likewise improves upon the OWMC in all categories. The only break from this pattern of improvement is the MRPE for the OTM call category for the GWMC, which is smaller for the GBM model than the Heston model which we can trace back to the excess volatility on the upside of the underlying for the geometric Brownian motion, which counters the exponential decay of the tails that we get from using the Kullback–Leibler divergence more aggressively.

**Table 6.** Intertemporal run results for the OWMC method and the GWMC method using geometric Brownian motion as an initial model for OTM puts ( $K < F$ ) and OTM calls ( $K > F$ ).

Method	Moneyness	MRPE	MAPE
OWMC	$K < F$	0.313	1.859
	$K > F$	0.267	1.163
	All	0.303	1.716
GWMC	$K < F$	0.151	0.899
	$K > F$	0.089	0.490
	All	0.136	0.816

**Table 7.** Intertemporal run results for the OWMC method and the GWMC method using the Heston model as an initial model for OTM puts ( $K < F$ ) and OTM calls ( $K > F$ ).

Method	Moneyness	MRPE	MAPE
OWMC	$K < F$	0.164	0.692
	$K > F$	0.202	0.696
	All	0.172	0.693
GWMC	$K < F$	0.074	0.401
	$K > F$	0.103	0.357
	All	0.080	0.392

## 6. Conclusions

We present a generalization of the weighted Monte Carlo calibration method for derivative pricing introduced by Avellaneda et al. [13]. The generalized WMC method is centered around relative entropy minimization similar to the original WMC method, but, instead of assuming a uniform prior measure for the sample paths, we compute a measure which distorts the probability of tail events. Furthermore, the proposed method partitions the set of benchmark instruments by maturity and assigns multiple weights per path, as opposed to the original method which assigns only one weight per path. Due to the specific nature of this twofold generalization, the original WMC naturally arises as a special case.

Through extensive numerical testing using S&P 500 options data, we show that the generalized method vastly outperforms the original WMC calibration method. The probability distortion and the increased flexibility yield a significant improvement in empirical fit to the whole range of options available on the underlying asset. This allows us to overcome the consistent implied volatility gap between the market data and what the original WMC method produces.

We also provide a detailed discussion on the utility maximization problem which is known to be dual to the relative entropy minimization formulation. The utility maximization formulation presented in this paper is technically speaking more general than the relative entropy minimization formulation, and it fits with any type of preferences which are of expected utility type. In future research, it would be worthwhile to study refinements of the weighted Monte Carlo method not only through more general probability distortion schemes but also through more general types of utility, such as recursive utility.

**Author Contributions:** Supervision, D.V.; writing—original draft, H.G.; and writing—review and editing, D.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Restrictions apply to the availability of these data. Option price data were obtained from IVolatility.com and are available at <https://www.ivolatility.com/>.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Abbreviations

The following abbreviations are used in this manuscript:

BFGS	Broyden–Fletcher–Goldfarb–Shanno
CPT	Cumulative Prospect Theory
GBM	geometric Brownian motion
GWMC	generalized weighted Monte Carlo
MAPE	mean average price error
MRPE	mean relative price error
OTM	out-of-the-money
OWMC	original weighted Monte Carlo
WMC	weighted Monte Carlo

### References

- Black, F.; Scholes, M. The pricing of options and corporate liabilities. *J. Political Econ.* **1973**, *81*, 637–654. [[CrossRef](#)]
- Rubinstein, M. Implied binomial trees. *J. Financ.* **1994**, *49*, 771–818. [[CrossRef](#)]
- Heston, S.L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.* **1993**, *6*, 327–343. [[CrossRef](#)]
- Derman, E.; Kani, I. Riding on a smile. *Risk* **1994**, *7*, 32–39.
- Dupire, B. Pricing with a smile. *Risk* **1994**, *7*, 18–20.
- Bates, D.S. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Rev. Financ. Stud.* **1996**, *9*, 69–107. [[CrossRef](#)]
- Andersen, L.; Andreasen, J. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Rev. Deriv. Res.* **2000**, *4*, 231–262. [[CrossRef](#)]
- Eraker, B.; Johannes, M.; Polson, N. The impact of jumps in volatility and returns. *J. Financ.* **2003**, *58*, 1269–1300. [[CrossRef](#)]
- Cont, R.; Tankov, P. Nonparametric calibration of jump-diffusion option pricing models. *J. Comput. Financ.* **2004**, *7*, 1–49. [[CrossRef](#)]
- Gilli, M.; Schumann, E. Calibrating option pricing models with heuristics. In *Natural Computing in Computational Finance*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 9–37.
- Björk, T. *Arbitrage Theory in Continuous Time*; Oxford University Press: Oxford, UK, 2009.
- Friedman, C.A.; Cao, W.; Huang, Y. Some Economically Meaningful Option Model Calibration Performance Measures. Working Paper, SSRN. 2014. Available online: <https://ssrn.com/abstract=2193803> (accessed on 16 December 2019).
- Avellaneda, M.; Buff, R.; Friedman, C.; Grandchamp, N.; Kruk, L.; Newman, J. Weighted Monte Carlo: A new technique for calibrating asset-pricing models. *Int. J. Theor. Appl. Financ.* **2001**, *4*, 91–119. [[CrossRef](#)]
- Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [[CrossRef](#)]
- Abbas, A.E.; Cadenbach, A.H.; Salimi, E. A Kullback–Leibler view of maximum entropy and maximum log-probability methods. *Entropy* **2017**, *19*, 232. [[CrossRef](#)]
- Guiasu, S.; Shenitzer, A. The principle of maximum entropy. *Math. Intell.* **1985**, *7*, 42–48. [[CrossRef](#)]
- Ludvigson, S.C. Advances in consumption-based asset pricing: Empirical tests. *Handb. Econ. Financ.* **2013**, *2*, 799–906.
- Carmona, R. *Indifference Pricing: Theory and Applications*; Princeton University Press: Princeton, NJ, USA, 2008.
- Tversky, A.; Kahneman, D. Advances in prospect theory: Cumulative representation of uncertainty. *J. Risk Uncertain.* **1992**, *5*, 297–323. [[CrossRef](#)]
- Camerer, C.F. Prospect Theory in the Wild: Evidence from the Field. In *Advances in Behavioral Economics*; Princeton University Press: Princeton, NJ, USA, 2011; pp. 148–161.
- Barberis, N. The psychology of tail events: Progress and challenges. *Am. Econ. Rev.* **2013**, *103*, 611–16. [[CrossRef](#)]
- Fox, C.R.; Rogers, B.A.; Tversky, A. Options traders exhibit subadditive decision weights. *J. Risk Uncertain.* **1996**, *13*, 5–17. [[CrossRef](#)]
- Polkovnichenko, V.; Zhao, F. Probability weighting functions implied in options prices. *J. Financ. Econ.* **2013**, *107*, 580–609. [[CrossRef](#)]
- Baele, L.; Driessen, J.; Londono, J.M.; Spalt, O.G. Cumulative Prospect Theory and the Variance Premium. Discussion Paper 12/2014-067, Netspar. 2014. Available online: <https://ssrn.com/abstract=2564498> (accessed on 16 December 2019).
- Vandenbroucke, J. A cumulative prospect view on portfolios that hold structured products. *J. Behav. Financ.* **2015**, *16*, 297–310. [[CrossRef](#)]

26. Friedman, C.A.; Cao, W.; Huang, Y.; Zhang, Y. Engineering More Effective Weighted Monte Carlo Option Pricing Models. Working Paper, SSRN. 2012. Available online: <https://ssrn.com/abstract=2193807> (accessed on 16 December 2019).
27. Broyden, C.G. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA J. Appl. Math.* **1970**, *6*, 76–90. [[CrossRef](#)]
28. Cochrane, J.H. *Asset Pricing: Revised Edition*; Princeton University Press: Princeton, NJ, USA, 2009.
29. Nau, R.F.; Jose, V.R.R.; Winkler, R.L. Duality between maximization of expected utility and minimization of relative entropy when probabilities are imprecise. In Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Application, Durham, UK, 14–18 July 2009; pp. 337–346.
30. Samperi, D.J. Inverse Problems, Model Selection and Entropy in Derivative Security Pricing. Ph.D. Thesis, New York University, New York, NY, USA, 1998.
31. Prelec, D. The probability weighting function. *Econometrica* **1998**, *66*, 497–527. [[CrossRef](#)]
32. Choquet, G. Theory of capacities. *Ann. L'institut Fourier* **1954**, *5*, 131–295. [[CrossRef](#)]
33. Elices, A.; Giménez, E. Weighted Monte Carlo. *Risk Mag.* **2006**, *19*, 78–83.