Data-Based Automatic Discretization of Nonparametric Distributions

Alexis Akira Toda

Accepted: 3 July 2020 / Published online: 20 July 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
Although using non-Gaussian distributions in economic models has become increasingly popular, currently there is no systematic way for calibrating a discrete distribution from the data without imposing parametric assumptions. This paper proposes a simple nonparametric calibration method based on the Golub-Welsch algorithm (Golub and Welsch in Math Comput 23(106): 221–230, 1969. https://doi.org/10.1090/S0025-5718-69-99647-1) for Gaussian quadrature. Applications to an asset pricing model and an optimal portfolio problem suggest that assuming normal instead of nonparametric shocks leads to up to 8% reduction in the equity premium and 17% overweighting in the stock portfolio because the investor underestimates the probability of crashes.

Keywords Calibration · Discrete approximation · Gaussian quadrature

1 Introduction

This paper studies the following problem, which applied theorists often encounter. A researcher would like to calibrate the parameters of a stochastic model. One of the model inputs is a probability distribution of shocks, which is to be approximated by a discrete distribution. Due to computational considerations, the researcher would like this distribution to have as few support points (nodes) as possible, say five. Given the data of shocks, how should the researcher calibrate the nodes and probabilities of this five-point distribution?

While there are many established methods for discretizing stochastic processes with Gaussian shocks such as Tauchen (1986), Tauchen and Hussey (1991), and Rouwenhorst (1995), discretizing non-Gaussian distributions remains relatively unexplored. While models with Gaussian shocks such as the long-run risk

1 See Farmer and Toda (2017) and the references therein for a detailed literature review.

Alexis Akira Toda
atoda@ucsd.edu

1 Department of Economics, University of California San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA
models (Bansal and Yaron, 2004) remain popular due to analytical tractability, it has become increasingly common in economics to study models with non-Gaussian shocks. For example, the rare disasters model (Rietz 1988; Barro 2006; Gabaix 2012) uses rare but large downward jumps to explain asset pricing puzzles. Toda and Walsh (2019) calibrate a three-point distribution from dividend growth data. One issue with discretizing non-Gaussian distributions is how to calibrate them. If we have a parametric density, it is possible to discretize it using the Gaussian quadrature as in Miller and Rice (1983) or the maximum entropy method as in Tanaka and Toda (2013, 2015) provided that we can compute some moments. However, it is not obvious how to obtain an $N$-point distribution that approximates the data well without imposing parametric assumptions. Because the degree of freedom of an $N$-point distribution is large ($2N - 1$, because there are $N$ nodes, $N$ probabilities, and the probabilities need to add up to 1), providing an automatic discretization method is valuable because it removes the arbitrariness of calibration.

Given the data, this paper proposes a simple method for automatically calibrating a discrete distribution with a specified number of grid points. The method is based on the observation that to compute the nodes and weights of the $N$-point Gaussian quadrature with some weighting function using the Golub and Welsch (1969) algorithm, one only needs to know the moments of the weighting function up to order $2N$. Therefore a natural way to discretize a nonparametric distribution is simply to feed the $2N$ sample moments into the Golub-Welsch algorithm. Since this method does not involve optimization (it is a matter of solving for the eigenvalues/vectors of a sparse $N \times N$ matrix), the implementation is easy and fast.\footnote{For the reader’s convenience, the MATLAB discretization code is posted at https://github.com/alexisakira/discretization; see NPGQ.m in the folder “Nonparametric Gaussian quadrature”.

As applications, I discretize the U.S. historical data on aggregate consumption growth and stock returns and solve an asset pricing model and an optimal portfolio problem with constant relative risk aversion utility. I consider two cases in which the investor uses the nonparametric and normal densities. I show that when the investor incorrectly believes that the consumption growth/stock returns distribution is lognormal, the equity premium shrinks by up to 8% and the stock portfolio is overweighted by up to 17% because he underestimates the probability of recessions/crashes. These examples show that the choice of the calibration method may matter quantitatively.

1.1 Related Literature

The closest paper to mine is Miller and Rice (1983), who use the Gaussian quadrature to discretize general parametric distributions. My method is different because it concerns the discretization of nonparametric distributions estimated from data. Tanaka and Toda (2013) consider the discretization of distributions on preassigned nodes by matching the moments using the maximum entropy principle, and Tanaka and Toda (2015) prove convergence and obtain an error estimate. Farmer and Toda...
Data-Based Automatic Discretization of Nonparametric…

(2017) consider the discretization of general non-Gaussian Markov processes by applying the Tanaka-Toda method to conditional distributions. In one of the applications, they discretize a nonparametric density on a preassigned grid by approximating it with a normal mixture. Since computing the nodes and weights of Gaussian quadrature does not require optimizing over parameters (unlike the maximum likelihood estimation of normal mixture parameters or solving the maximum entropy problem), my method is easier and faster to implement, and the grid is chosen endogenously. On the other hand, the Farmer and Toda (2017) method can discretize general Markov processes, whereas the proposed method in this paper is designed to discretize a single distribution.

2 Discretizing a Nonparametric Density

Suppose for the moment that the nonparametric density \( f(x) \) is known. Since stochastic models often involve expectations, we would like to find nodes \( \{x_n\}_{n=1}^N \) and weights \( \{w_n\}_{n=1}^N \) such that

\[
E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x) \, dx \approx \sum_{n=1}^{N} w_n g(x_n),
\]

(1)

where \( g \) is a general integrand and \( X \) is a random variable with density \( f(x) \). The right-hand side of (1) defines an \( N \)-point quadrature formula.

When (1) is exact (i.e., \( \approx \) becomes =) for all polynomials of degree up to \( D \), we say that the quadrature formula has degree of exactness \( D \). Since the degree of freedom in an \( N \)-point quadrature formula is \( 2N \) (because there are \( N \) nodes and \( N \) weights), we cannot expect to integrate more than \( 2N \) monomials \( g(x) = 1, x, \ldots, x^{2N-1} \) exactly. When the quadrature formula (1) is exact for these monomials, or equivalently when it has degree of exactness \( 2N-1 \), we call the formula Gaussian. The following Golub-Welsch algorithm provides an efficient way to compute the nodes and weights of the Gaussian quadrature (“Appendix C” provides more theoretical background.)

Algorithm 1 (Golub and Welsch 1969)
1. Select a number of quadrature nodes \( N \in \mathbb{N} \).
2. For \( k = 0, 1, \ldots, 2N \), compute the \( k \)-th moment of the density \( m_k = \int x^k f(x) \, dx \).
3. Define the matrix of moments \( M = (M_{ij})_{1 \leq i,j \leq N+1} \) by \( M_{ij} = m_{i+j-2} \).
4. Compute the Cholesky factorization \( M = R^T R \). Let \( R = (r_{ij})_{1 \leq i,j \leq N+1} \).
5. Define \( a_1 = r_{12}/r_{11} \), \( a_n = r_{n+1,n} r_{n,n-1}/r_{n,n} \) (\( n = 2, \ldots, N \)), and \( \beta_n = r_{n+1,n}/r_{n,n} \) (\( n = 1, \ldots, N-1 \)). Define the \( N \times N \) symmetric tridiagonal matrix

\[
T_N = \begin{bmatrix}
\alpha_1 & \beta_1 & 0 & \cdots & 0 \\
\beta_1 & \alpha_2 & \beta_2 & \ddots & \vdots \\
0 & \beta_2 & \alpha_3 & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & \beta_{N-1} \\
0 & \cdots & 0 & \beta_{N-1} & \alpha_N
\end{bmatrix}
\]

(2)
6. Compute the eigenvalues \( \{ x_n \}_{n=1}^N \) of \( T_N \) and the corresponding eigenvectors \( \{ v_n \}_{n=1}^N \). \( x_n \) and \( v_n \) are the nodes of the Gaussian quadrature and the weights \( \{ w_n \}_{n=1}^N \) are given by \( w_n = m_0 v_n^2 / \| v_n \|^2 > 0 \), where \( v_n = (v_{n1}, \ldots, v_{nm}) \).

Once we compute the nodes \( \{ x_n \}_{n=1}^N \) and weights \( \{ w_n \}_{n=1}^N \), we can use them as the discrete approximation of the density \( f \).

Note that the only inputs to the Golub-Welsch algorithm 1 are the number of nodes \( N \) and the moments \( m_k = \int x^k f(x) \, dx \) of the density \( f \), where \( k = 0, \ldots, 2N \). Therefore a natural idea for discretizing a nonparametric density given the data \( \{ x_i \}_{i=1}^I \) is simply to feed the sample moments into the Golub-Welsch algorithm 1. Summarizing the above observations, we obtain the following algorithm for the data-based automatic discretization of nonparametric distributions.

Algorithm 2 (Automatic discretization of nonparametric distributions)

1. Given the data \( \{ x_i \}_{i=1}^I \) and the desired number of discrete points \( N \), for \( k = 0, \ldots, 2N \) compute the \( k \)-th sample moment
   \[
   \hat{m}_k = \frac{1}{I} \sum_{i=1}^I x_i^k. \tag{3}
   \]
2. Feed these moments \( \{ \hat{m}_k \}_{k=0}^{2N} \) into the Golub-Welsch algorithm 1 to compute the nodes \( \{ \hat{x}_n \}_{n=1}^N \) and weights \( \{ w_n \}_{n=1}^N \). The desired discretization assigns probability \( w_n \) on the point \( \hat{x}_n \).

Because the \( N \)-point Gaussian quadrature has degree of exactness \( 2N - 1 \), by construction the \( N \)-point discretization matches the sample moments of data up to order \( 2N - 1 \) (and up to numerical error). Since by the Gauss-Markov theorem the sample moment (3) is the best linear unbiased estimator (BLUE) of the population moment, that is, \( \hat{m}_k \) has the minimum mean-squared error among all estimates of the form \( \sum_{i=1}^I a_i x_i^k \), where \( \{ a_i \}_{i=1}^I \) are some weights, Algorithm 2 is in a sense optimal. We can also justify Algorithm 2 based on the properties of the moment generating function. Curties (1942, Theorem 1) shows that, for random variables having finite moment generating functions in the neighborhood of 0, there is a one-to-one correspondence between the distribution and its moment generating function. Therefore by applying the Taylor approximation, we expect that we can approximate the properties of a random variable by matching sufficiently many moments. “Appendix B” shows that the accuracy of the proposed method exceeds that of using a parametric distribution when the latter is misspecified.

As a practical consideration, to improve the numerical stability, it is desirable to standardize the data before feeding into the algorithm. Therefore given the data \( \{ x_i \}_{i=1}^I \), one should compute the sample mean and variance
\[
\mu = \frac{1}{I} \sum_{i=1}^I x_i \quad \text{and} \quad \sigma^2 = \frac{1}{I} \sum_{i=1}^I (x_i - \mu)^2,
\]
define the standardized data \( z_i := (x_i - \mu) / \sigma \), compute the moments of standardized data \( \hat{m}_k = \frac{1}{I} \sum_{i=1}^I z_i^k \), and use the Golub-Welsch algorithm to obtain the nodes \( \{ \hat{z}_n \}_{n=1}^N \) and weights \( \{ w_n \}_{n=1}^N \). Then the nonparametric discretization of the original data is given by \( \hat{x}_n = \mu + \sigma \hat{z}_n \) with the same weights.
3 Applications

In this section, I illustrate the usefulness of the proposed method using minimal economic examples. In the first application, I solve a consumption-based asset pricing model by calibrating the consumption growth distribution from the data. In the second application, I solve a static optimal portfolio problem by calibrating the stock returns distribution from the data. In each case I compare the nonparametric solution to the case when the distribution is assumed to be lognormal and show that the choice of the discretization method matters quantitatively.

3.1 Asset Pricing Model

Consider a plain-vanilla Lucas (1978) consumption-based asset pricing model, where a representative agent with constant relative risk aversion (CRRA) utility consumes the aggregate endowment. Endowment growth is assumed to be independent and identically distributed (i.i.d.) over time with a general distribution. A stock is modeled as a claim to the aggregate endowment, and the net supply of the risk-free asset is zero.

Let \( R > 0 \) be the gross stock return (which is an i.i.d. random variable) and \( R_f > 0 \) be the gross risk-free rate. Letting \( \gamma > 0 \) be the relative risk aversion coefficient of the representative agent, the log equity premium is given by

\[
E[\log R] - \log R_f = \log M(1) + \log M(-\gamma) - \log M(1 - \gamma) > 0, \tag{2}
\]

where \( M(s) = E[e^{sx}] \) is the moment generating function of the log endowment growth \( x_t = \log(e_t/e_{t-1}) \).

I obtain the annual data on U.S. real per capita consumption expenditures for the period 1889–2018 and apply Algorithm 2 to the log consumption growth \( \log C_t/C_{t-1} \) to obtain a discrete distribution with nodes \( \{\bar{x}_n\}_{n=1}^N \) and weights \( \{w_n\}_{n=1}^N \), where I choose the number of points \( N = 5 \) (increasing the number of points further does not change the results). I then compute the moment generating function as \( M(s) = \sum_{n=1}^N w_n e^{sx_n} \) and substitute into (2) for \( s = 1, -\gamma, 1 - \gamma \) to compute the log equity premium for the range \( \gamma \in [1, 10] \).

Figure 1a shows the histogram of the log consumption growth distribution as well as the nonparametric kernel density estimator and the normal distribution fitted by maximum likelihood. We can see that the histogram and the nonparametric density

---

3 See the appendix for the derivation. With i.i.d. endowment growth, it is easy to show that the log equity premium depends only on the risk aversion and it does not matter whether the utility is additive or recursive (Epstein-Zin).

4 For this purpose I combine two data sets. The first is the spreadsheet of Robert Shiller (http://www.econ.yale.edu/~shiller/data/chapt26.xlsx), which contains data up to 2009. The second is the real consumption expenditures and population data from FRED (https://fred.stlouisfed.org), which is available for 1929–2018. I compute the real per capita consumption expenditures from the FRED data by dividing consumption by population. Normalizing the Shiller and FRED series to 1 in year 2000, the two series are nearly identical in overlapping years (1929–2009). Therefore I use the Shiller data up to 2000 and FRED data from 2001 on.
have slightly longer tail (expansions and recessions) than the normal distribution. Figure 1b shows the log equity premium (2) for the two models. Fig. 1c shows the relative increase in the equity premium when using the nonparametric distribution. Using the nonparametric distribution, the equity premium increases by 8% relative to the normal distribution when the relative risk aversion is $\gamma = 10$. Thus using nonparametric distributions help to explain asset pricing puzzles.

### 3.2 Optimal Portfolio Problem

Consider a CRRA investor with relative risk aversion $\gamma > 0$. Letting $R > 0$ be the gross stock return, $R_f > 0$ be the gross risk-free rate, and $\theta$ be the fraction of wealth (portfolio share) invested in the stock, the investor’s optimal portfolio problem is

---

5 To make the results comparable, I discretize the normal density using the Golub-Welsch algorithm with the corresponding normal density, which is equivalent to the well-known Gauss-Hermite quadrature.
I obtain the annual data on U.S. nominal stock returns, risk-free rate, and inflation for the period 1927–2016 from the spreadsheet of Amit Goyal. For the stock returns I use the CRSP volume-weighted index including dividends. I convert these returns into real log returns and calibrate the log risk free rate as the sample average. The result is $R_f = 1.0045$. I then apply Algorithm 2 to the log excess returns $\log R - \log R_f$ to obtain a discrete distribution with nodes $\{x_n\}_{n=1}^N$ and weights $\{w_n\}_{n=1}^N$, where I choose the number of points $N = 5$ (increasing the number of points further does not change the results). The gross stock return in state $n$ is defined by $R_n = R_f e^{x_n}$, which occurs with probability $w_n$. Finally, I numerically solve the optimal portfolio problem (3). Figure 2 shows the results when we change the relative risk aversion in the range $\gamma \in [1, 7]$.

$$\max_\theta \frac{1}{1-\gamma} E[(R \theta + R_f (1-\theta))^{1-\gamma}]. \quad (3)$$

Fig. 2 Excess returns distribution and numerical solution

---

6 The spreadsheet is at http://www.hec.unil.ch/agoyal/docs/Predic torData2016.xlsx. Using monthly or quarterly data give qualitatively similar results, though slightly less extreme quantitatively.
Figure 2a shows the histogram of the log excess returns distribution as well as the nonparametric kernel density estimator and the normal distribution fitted by maximum likelihood. We can see that the histogram and the nonparametric density have long left tails corresponding to stock market crashes, which the normal distribution misses. Figure 2b shows the optimal portfolio $\theta$ for the two models. We can see that when the investor incorrectly believes that the stock returns distribution is log-normal, he overweights the stock portfolio because he underestimates the probability of crashes. Figure 2c shows the percentage of this overweight (portfolio error) $\theta_N/\theta_{NP} - 1$, where N and NP stand for normal and nonparametric densities. The portfolio error is substantial, in the range of 4–17%.

4 Concluding Remarks

This paper has proposed a simple, automatic method for discretizing a nonparametric distribution, given the data. Using an asset pricing model and an optimal portfolio problem as a laboratory, I showed that the error from using a parametric distribution (such as the normal distribution) can be substantial.

A natural extension is to consider the discretization of Markov processes with nonparametric shocks. For example, one may be tempted to apply the kernel density estimation and Gaussian quadrature in the Tauchen and Hussey (1991) method to discretize the AR(1) process

$$x_t = \rho x_{t-1} + \varepsilon_t,$$

where $|\rho| < 1$ and the innovations $\{\varepsilon_t\}_{t=0}^{\infty}$ are independent and identically distributed according to some probability density function $f$. However, it is well known that the Tauchen-Hussey method is not accurate when the persistence $\rho$ is moderately high (Flodén 2008). Using the AR(1) asset pricing model in the Online Appendix of Farmer and Toda (2017) to evaluate the solution accuracy, I found that the Gaussian quadrature-based methods for discretizing Markov processes is even less accurate when the shock distribution is nonparametric. Therefore for such processes, it is preferable to use the Farmer and Toda (2017) maximum entropy method with an even-spaced grid (see their Sect. 4.3.3 for an example).

Finally, although I proposed my method as a tool for discretization, it can also be used as a quadrature method. For example, Pohl et al. (2018) solve the Bansal and Yaron (2004) long run risks model using the projection method and Gauss-Hermite quadrature, but that is because the model is assumed to have normal shocks. If instead the researcher wishes to use nonparametric shocks, my method can be directly used to construct a quadrature rule from data.
Appendix A: Solving the Asset Pricing Model

In this appendix, I compute the risk-free rate, expected stock returns, and equity premium in a representative-agent asset pricing model when the agent has discount factor $\beta$, relative risk aversion $\gamma$, and consumption growth is i.i.d. Since the stochastic discount factor is $\beta(C_{t+1}/C_t)^{-\gamma}$, the gross risk-free rate $R_f$ satisfies

$$\frac{1}{R_f} = E[\beta(C_{t+1}/C_t)^{-\gamma}] = \beta E[e^{-\gamma \Delta \log C}] = \beta M(-\gamma) \iff R_f = \frac{1}{\beta M(-\gamma)},$$

where $M(s) = E[e^{s \Delta \log C}]$ is the moment generating function of log consumption growth.\(^7\)

The asset pricing equation for the stock (consumption claim) is

$$P_t = E[\beta(C_{t+1}/C_t)^{-\gamma}(P_{t+1} + C_{t+1})].$$

dividing both sides by $C_t$ and letting $v = P_t/C_t$ be the price-dividend ratio (which must be constant in an i.i.d. environment), we obtain

$$v = E[\beta(C_{t+1}/C_t)^{1-\gamma}(v + 1)] = \beta M(1 - \gamma)(v + 1) \iff v = \frac{\beta M(1 - \gamma)}{1 - \beta M(1 - \gamma)}.$$

Since the gross stock return is

$$R_{t+1} = \frac{P_{t+1} + C_{t+1}}{P_t} = \frac{C_{t+1}}{C_t} \frac{P_{t+1}/C_{t+1} + 1}{P_t/C_t} = \frac{C_{t+1}}{C_t} \frac{1}{\beta M(1 - \gamma)},$$

taking expectations, the expected stock return is

$$E[R] = \frac{M(1)}{\beta M(1 - \gamma)}.$$

Therefore the log equity premium is

$$\log E[R] - \log R_f = \log \frac{M(1)M(-\gamma)}{M(1 - \gamma)}.$$

Intuitively, the equity premium should be nonnegative. Indeed, we can prove that that is the case.

**Proposition 1** Let $M$ be a moment generating function. Then

$$\log M(1) + \log M(-\gamma) \geq \log M(1 - \gamma).$$

**Proof** Let $X$ be log consumption growth and take any $s_1, s_2$ and $t \in (0, 1)$. Letting $p = \frac{1}{1-t}$ and $q = \frac{1}{t}$, we have $1/p + 1/q = 1$. By Hölder’s inequality, we obtain

\(^7\) The usefulness of using the moment generating function for solving asset pricing models was observed by Tsionas (2003).
Taking the logarithm, we obtain
\[
\log M((1 - t)s_1 + ts_2) \leq (1 - t) \log M(s_1) + t \log M(s_2),
\]
so \(\log M\) is convex (\(M\) is log-convex). For notational simplicity, let \(f(s) = \log M(s)\).
Since \(\gamma > 0\) we have \(-\gamma < 0, 1 - \gamma < 1\). Since
\[
0 = (1 - t)(-\gamma) + t \iff t = \frac{\gamma}{\gamma + 1},
\]
\[
1 - \gamma = (1 - t)(-\gamma) + t \iff t = \frac{1}{\gamma + 1},
\]
By the convexity of \(f\) we obtain
\[
\begin{align*}
    f(0) &\leq \frac{1}{\gamma + 1} f(-\gamma) + \frac{\gamma}{\gamma + 1} f(1), \\
    f(1 - \gamma) &\leq \frac{\gamma}{\gamma + 1} f(-\gamma) + \frac{1}{\gamma + 1} f(1).
\end{align*}
\]
Adding these two inequalities and noting that \(f(0) = \log M(0) = 0\) because \(M(0) = 1\), it follows that \(f(1 - \gamma) \leq f(-\gamma) + f(1)\), which is the conclusion. \(\square\)

**Appendix B: Accuracy**

As in any numerical method, evaluating the accuracy is very important. In this section, I evaluate the accuracy of the proposed method using the optimal portfolio problem in Sect. 3.2 as a laboratory. I consider two specifications. Because the proposed method is designed to discretize nonparametric (in particular, non-normal) distributions, in the first experiment I suppose that the excess returns distribution is a normal mixture. In the second experiment, I suppose that the excess returns distribution is normal to see how worse the nonparametric Gaussian quadrature method performs relative to correctly specified parametric methods.

**Appendix B.1: Normal Mixture Distribution**

I design the numerical experiment as follows. First I fit a normal mixture distribution with two components to the annual log excess returns data. The proportion, mean, and standard deviation of each mixture components are \(p = (p_j) = (0.1392, 0.8608)\), \(\mu = (\mu_j) = (-0.2242, 0.1064)\), and
\( \sigma = (\sigma_j) = (0.2164, 0.1453) \), respectively. Next, I assume that the true excess returns distribution is this normal mixture and solve the optimal portfolio problem for relative risk aversion \( \gamma \in \{2, 4, 6\} \) using the Gaussian quadrature (Golub-Welsch algorithm 1) for normal mixtures with 11 points. Finally, I generate random numbers from this normal mixture with various sample sizes, discretize these distributions with various methods, and compute the optimal portfolio. I repeat this procedure with \( M = 1,000 \) Monte Carlo replications and compute the relative bias and mean absolute error (MAE) defined by

\[
\text{Bias} = \frac{1}{M} \sum_{m=1}^{M} \left( \hat{\theta}_m / \theta^* - 1 \right),
\]

\[
\text{MAE} = \frac{1}{M} \sum_{m=1}^{M} \left| \hat{\theta}_m / \theta^* - 1 \right|
\]

where \( \hat{\theta}_m \) is the optimal portfolio from simulation \( m \) and \( \theta^* \) is the theoretical optimal portfolio. For the sample size I consider \( T = 100, 1,000, 10,000 \), and for the number of quadrature nodes I consider \( N = 3, 5, 7, 9 \) (increasing \( N \) further does not change the results). For the discretization method I consider three cases. The first is the nonparametric Gaussian quadrature method (Algorithm 2), which I refer to as “NP-GQ”. The second is the Gauss-Hermite quadrature, where the mean and standard deviation are estimated by maximum likelihood. This is the most natural method if the returns distribution is lognormal. The third is the maximum entropy method proposed by Tanaka and Toda (2013, 2015) and Farmer and Toda (2017) where the kernel density estimator (with Gaussian kernel) is fed into, which I refer to as “NP-ME”. For this method one needs to assign the grid and the number of moments to match. Following Corollary 3.5 of Farmer and Toda (2017), I use an even-spaced grid centered at the sample mean that spans \( \sqrt{2(N - 1)} \) times the sample standard deviation at both sides, where \( N \) is the number of grid points. I match 4 sample moments whenever possible for \( N \geq 5 \), and otherwise I match 2 sample moments (mean and variance). For more details on the exact algorithm, please refer to Tanaka and Toda (2013, 2015) and Sects. 2 and 3.2 of Farmer and Toda (2017).

Tables 1 and 2 show the relative bias and mean absolute error of the optimal portfolio, respectively. As expected, the optimal portfolio computed using Gauss-Hermite is biased upwards because it uses the normal distribution, which underestimates the probability of crashes. Among the two nonparametric discretization methods, NP-GQ uniformly outperforms NP-ME in terms of bias and mean absolute error, especially when the sample size is small (\( T = 100 \)). For \( N = 3 \) grid points, in which case it is impossible to match 4 moments with NP-ME, the proposed NP-GQ method performs significantly better, probably because NP-GQ matches \( 2 \times 3 - 1 = 5 \) moments. Finally, increasing \( N \) beyond 5 does not improve the bias or the mean absolute error for NP-GQ, which suggests that using a five-point distribution is enough (at least for solving this portfolio problem).
Appendix B.2: Normal Distribution

I conduct a similar experiment assuming that the log excess returns distribution is normal. For this exercise, I compare the nonparametric Gaussian quadrature method...
Data-Based Automatic Discretization of Nonparametric…

Table 3 Relative bias of the optimal portfolio (normal)

| Method   | NP-GQ | Gauss-Hermite | Tauchen |
|----------|-------|---------------|---------|
| $T$  | $N$  | $\gamma = 2$ | $\gamma = 4$ | $\gamma = 6$ | $\gamma = 2$ | $\gamma = 4$ | $\gamma = 6$ | $\gamma = 2$ | $\gamma = 4$ | $\gamma = 6$ |
| 100     | 3     | 0.033         | 0.030      | 0.029     | 0.026         | 0.027      | 0.027     | -0.083     | -0.084      | -0.085       |
| 5       | 0.030 | 0.030         | 0.029     | 0.021     | 0.021         | 0.027      | 0.027     | -0.083     | -0.084      | -0.085       |
| 7       | 0.030 | 0.030         | 0.029     | 0.021     | 0.021         | 0.027      | 0.027     | -0.083     | -0.084      | -0.085       |
| 9       | 0.030 | 0.030         | 0.029     | 0.021     | 0.021         | 0.027      | 0.027     | -0.083     | -0.084      | -0.085       |
| 1,000   | 3     | 0.006         | 0.005      | 0.005     | 0.005         | 0.005      | 0.005     | -0.100     | -0.103      | -0.104       |
| 5       | 0.005 | 0.005         | 0.005     | 0.004     | 0.005         | 0.005      | 0.005     | -0.100     | -0.103      | -0.104       |
| 7       | 0.005 | 0.005         | 0.005     | 0.004     | 0.005         | 0.005      | 0.005     | -0.100     | -0.103      | -0.104       |
| 9       | 0.005 | 0.005         | 0.005     | 0.004     | 0.005         | 0.005      | 0.005     | -0.100     | -0.103      | -0.104       |
| 10,000  | 3     | 0.000         | -0.001    | 0.000     | 0.000         | 0.000      | 0.000     | -0.104     | -0.108      | -0.109       |
| 5       | 0.000 | 0.000         | 0.000     | 0.000     | 0.000         | 0.000      | 0.000     | -0.104     | -0.108      | -0.109       |
| 7       | 0.000 | 0.000         | 0.000     | 0.000     | 0.000         | 0.000      | 0.000     | -0.104     | -0.108      | -0.109       |
| 9       | 0.000 | 0.000         | 0.000     | 0.000     | 0.000         | 0.000      | 0.000     | -0.104     | -0.108      | -0.109       |

The table reports the relative bias of the optimal portfolio defined by (4a) with a normal excess returns distribution. See Table 1 for the definition of variables.

Table 4 Relative mean absolute error of the optimal portfolio (normal)

| Method   | NP-GQ | Gauss-Hermite | Tauchen |
|----------|-------|---------------|---------|
| $T$  | $N$  | $\gamma = 2$ | $\gamma = 4$ | $\gamma = 6$ | $\gamma = 2$ | $\gamma = 4$ | $\gamma = 6$ | $\gamma = 2$ | $\gamma = 4$ | $\gamma = 6$ |
| 100     | 3     | 0.224         | 0.224      | 0.224     | 0.217         | 0.221      | 0.221     | 0.199       | 0.205       | 0.205       |
| 5       | 0.220 | 0.224         | 0.224     | 0.212     | 0.221         | 0.222      | 0.222     | 0.199       | 0.205       | 0.205       |
| 7       | 0.220 | 0.224         | 0.224     | 0.212     | 0.221         | 0.222      | 0.222     | 0.199       | 0.205       | 0.205       |
| 9       | 0.220 | 0.224         | 0.224     | 0.212     | 0.221         | 0.222      | 0.222     | 0.199       | 0.205       | 0.205       |
| 1,000   | 3     | 0.069         | 0.069      | 0.069     | 0.068         | 0.069      | 0.069     | 0.106       | 0.110       | 0.111       |
| 5       | 0.068 | 0.070         | 0.070     | 0.067     | 0.069         | 0.069      | 0.069     | 0.106       | 0.110       | 0.111       |
| 7       | 0.068 | 0.070         | 0.070     | 0.067     | 0.069         | 0.069      | 0.069     | 0.106       | 0.110       | 0.111       |
| 9       | 0.068 | 0.070         | 0.070     | 0.067     | 0.069         | 0.069      | 0.069     | 0.106       | 0.110       | 0.111       |
| 10,000  | 3     | 0.021         | 0.021      | 0.021     | 0.021         | 0.021      | 0.021     | 0.104       | 0.108       | 0.109       |
| 5       | 0.021 | 0.021         | 0.021     | 0.020     | 0.021         | 0.021      | 0.021     | 0.104       | 0.108       | 0.109       |
| 7       | 0.021 | 0.021         | 0.021     | 0.020     | 0.021         | 0.021      | 0.021     | 0.104       | 0.108       | 0.109       |
| 9       | 0.021 | 0.021         | 0.021     | 0.020     | 0.021         | 0.021      | 0.021     | 0.104       | 0.108       | 0.109       |

The table reports the relative mean absolute error of the optimal portfolio defined by (4b) with a normal excess returns distribution. See Table 1 for the definition of variables.

to Gauss-Hermite quadrature (Tauchen and Hussey 1991) and the Tauchen (1986) method. For the Tauchen method, I use the MATLAB code written by Martin Flodén.

Tables 3 and 4 show the relative bias and mean absolute error of the optimal portfolio, respectively. The Tauchen method, which is popular in applied works, is very poor, with bias and mean absolute errors of about 10–20%. Since the poor
performance of the Tauchen method is documented elsewhere, for example Flodén (June 2008) and Farmer and Toda (2017), this result is not surprising. It is another evidence that more sophisticated methods such as Farmer and Toda (2017) should be used. What is surprising, however, is that the nonparametric Gaussian mixture method (NP-GQ), which is fully nonparametric, performs similarly to (only slightly worse than) Gauss-Hermite quadrature, which is the most natural method when the distribution is normal. This experiment suggests that even if the data generating process is correctly specified, using the nonparametric Gaussian mixture does not necessarily compromise accuracy.

Online Appendix (not for publication)

Appendix C: Gaussian Quadrature

In this appendix we prove some properties of the Gaussian quadrature. Let \( w(x) \) be a weighting function and assume that \( \int_a^b w(x)x^n \, dx \) exists for all \( n \geq 0 \), where \(-\infty \leq a < b \leq \infty \) are fixed. For functions \( f, g \), define the inner product \((f, g)\) by

\[(f, g) = \int_a^b w(x)f(x)g(x) \, dx. \tag{5}\]

As usual, define the norm of \( f \) by \( \|f\| = \sqrt{(f, f)} \). For notational simplicity, let us omit \( a, b \), so \( \int \) means \( \int_a^b \).

The first step is to construct orthogonal polynomials \( \{p_n(x)\}_{n=0}^N \) corresponding to the inner product (5).

**Definition 1 (Orthogonal polynomial)** The polynomials \( \{p_n(x)\}_{n=0}^N \) are called **orthogonal** if

(i) \( \deg p_n = n \) and the leading coefficient of \( p_n \) is 1, and

(ii) for all \( m \neq n \), we have \( (p_m, p_n) = 0 \).

Some authors require that the polynomials are orthonormal, so \( (p_n, p_n) = 1 \). In this paper we normalize the polynomials by requiring that the leading coefficient is 1, which is useful for computation. The following three-term recurrence relation (TTRR) shows the existence of orthogonal polynomials and provides an explicit algorithm for computing them.

**Proposition 2 (Three-term recurrence relation, TTRR)** Let \( p_0(x) = 1 \), \( p_1(x) = x - \frac{(x, p_0)}{\|p_0\|^2} \), and for \( n \geq 1 \) define

\[ p_{n+1}(x) = \left( x - \frac{(xp_n, p_n)}{\|p_n\|^2} \right) p_n(x) - \frac{\|p_n\|^2}{\|p_{n-1}\|^2} p_{n-1}(x). \tag{6}\]

Then \( p_n(x) \) is the degree \( n \) orthogonal polynomial.
Proof Let us show by induction on \( n \) that

(i) \( p_n \) is a degree \( n \) polynomial with leading coefficient \( 1 \), and

(ii) \( (p_n, p_m) = 0 \) for all \( m < n \).

The claim is trivial for \( n = 0 \). For \( n = 1 \), by construction \( p_1 \) is a degree 1 polynomial with leading coefficient \( 1 \), and since \( p_0(x) = 1 \), we obtain

\[
(p_1, p_0) = \left( x - \frac{(xp_0, p_0)}{\|p_0\|^2} \right) p_0 = (xp_0, p_0) - (xp_0, p_0) = 0.
\]

Suppose the claim holds up to \( n \). Then for \( n + 1 \), by (6) the leading coefficient of \( p_{n+1} \) is the same as that of \( xp_n \), which is 1. If \( m = n \), then

\[
(p_{n+1}, p_n) = \left( x - \frac{(xp_n, p_n)}{\|p_n\|^2} \right) p_n = \frac{\|p_n\|^2}{\|p_{n-1}\|^2} (p_{n-1}, p_n)
\]

\[
= (xp_n, p_n) - (xp_n, p_n) = \frac{\|p_n\|^2}{\|p_{n-1}\|^2} (p_{n-1}, p_n) = 0.
\]

If \( m = n - 1 \), then

\[
(p_{n+1}, p_{n-1}) = \left( x - \frac{(xp_n, p_n)}{\|p_n\|^2} \right) p_n - \frac{\|p_n\|^2}{\|p_{n-1}\|^2} (p_{n-1}, p_{n-1})
\]

\[
= (xp_n, p_{n-1}) - \frac{(xp_n, p_n)}{\|p_n\|^2} (p_{n-1}, p_{n-1}) - \frac{\|p_n\|^2}{\|p_{n-1}\|^2} p_n
\]

\[
= (p_n, xp_{n-1}) - \|p_n\|^2.
\]

Since the leading coefficients of \( p_n \) are 1, we can write \( xp_{n-1}(x) = p_n(x) + q(x) \), where \( q(x) \) is a polynomial of degree at most \( n - 1 \). Clearly \( q \) can be expressed as a linear combination of \( p_0, p_1, \ldots, p_{n-1} \), so \( (p_n, q) = 0 \). Therefore

\[
(p_{n+1}, p_{n-1}) = (p_n, p_n + q) - \|p_n\|^2 = \|p_n\|^2 + (p_n, q) - \|p_n\|^2 = 0.
\]

Finally, if \( m < n - 1 \), then

\[
(p_{n+1}, p_m) = \left( x - \frac{(xp_m, p_m)}{\|p_m\|^2} \right) p_m - \frac{\|p_m\|^2}{\|p_{n-1}\|^2} (p_{n-1}, p_m)
\]

\[
= (xp_m, p_m) - \frac{(xp_m, p_m)}{\|p_m\|^2} (p_{n-1}, p_m) - \frac{\|p_m\|^2}{\|p_{n-1}\|^2} (p_{n-1}, p_m)
\]

\[
= (p_n, xp_m) = 0
\]

because \( xp_m \) is a polynomial of degree \( 1 + m < n \). \( \square \)
The following lemma shows that an degree $n$ orthogonal polynomial has exactly $n$ real roots (so they are all simple).

**Lemma 1** $p_n(x)$ has exactly $n$ real roots on $(a, b)$.

**Proof** By the fundamental theorem of algebra, $p_n(x)$ has exactly $n$ roots in $\mathbb{C}$. Suppose on the contrary that $p_n(x)$ has less than $n$ real roots on $(a, b)$. Let $x_1, \ldots, x_k$ ($k < n$) those roots at which $p_n(x)$ changes its sign. Let $q(x) = (x - x_1) \cdots (x - x_k)$. Since $p_n(x)q(x) > 0$ (or $< 0$) almost everywhere on $(a, b)$, we have

$$(p_n, q) = \int w(x)p_n(x)q(x) \, dx \neq 0.$$ 

On the other hand, since deg $q = k < n$, we have $(p_n, q) = 0$, which is a contradiction. $\square$

The following theorem shows that using the $N$ roots of the degree $N$ orthogonal polynomial $p_N(x)$ as quadrature nodes and choosing specific weights, we can integrate all polynomials of degree up to $2N - 1$ exactly. Thus Gaussian quadrature always exists.

**Theorem 1** (Gaussian quadrature) Let $a < x_1 < \cdots < x_N < b$ be the $N$ roots of the degree $N$ orthogonal polynomial $p_N$ and define

$$w_n = \int w(x)L_n(x) \, dx$$

for $n = 1, \ldots, N$, where

$$L_n(x) = \prod_{m \neq n} \frac{x - x_m}{x_n - x_m}$$

is the degree $N - 1$ polynomial that takes value 1 at $x_n$ and 0 at $x_m$ ($m \in \{1, \ldots, N\} \setminus n$). Then

$$\int w(x)p(x) \, dx = \sum_{n=1}^N w_np(x_n)$$ (7)

for all polynomials $p(x)$ of degree up to $2N - 1$.

**Proof** Since deg $p \leq 2N - 1$ and deg $p_N = N$, we can write

$$p(x) = p_N(x)q(x) + r(x),$$

where deg $q, \text{deg } r \leq N - 1$. Since $q$ can be expressed as a linear combination of orthogonal polynomials of degree up to $N - 1$, we have $(p_N, q) = 0$. Hence
\[ \int w(x)p(x) \, dx = \langle p_N, q \rangle + \int w(x)r(x) \, dx = \int w(x)r(x) \, dx. \]

On the other hand, since \( \{x_n\}_{n=1}^N \) are roots of \( p_N \), we have
\[ p(x_n) = p_N(x_n)q(x_n) + r(x_n) = r(x_n) \]
for all \( n \), so in particular
\[ \sum_{n=1}^N w_n p(x_n) = \sum_{n=1}^N w_n r(x_n). \]

Therefore it suffices to show (7) for polynomials \( r \) of degree up to \( N - 1 \). Let us show that
\[ r(x) = \sum_{n=1}^N r(x_n)L_n(x) \]
identically. To see this, let \( \tilde{r} \) be the right-hand side. Since \( L_n(x_m) = \delta_{mn} \) (Kronecker’s delta), we have
\[ \tilde{r}(x_m) = \sum_{n=1}^N r(x_n)L_n(x_m) = \sum_{n=1}^N \delta_{mn}r(x_n) = r(x_m), \]
so \( r \) and \( \tilde{r} \) agree on \( N \) distinct points \( \{x_n\}_{n=1}^N \). Since each \( L_n(x) \) is a degree \( N - 1 \) polynomial, we have \( \deg \tilde{r} \leq N - 1 \). Therefore it must be \( r = \tilde{r} \).

Since \( r \) can be represented as a linear combination of \( L_n \)'s, it suffices to show (7) for all \( L_n \)'s. But since by definition
\[ \int w(x)L_n(x) \, dx = w_n = \sum_{m=1}^N w_m \delta_{mn} = \sum_{m=1}^N w_m L_n(x_m), \]
the claim is true.

\[ \square \]

In practice, how can we compute the nodes \( \{x_n\}_{n=1}^N \) and weights \( \{w_n\}_{n=1}^N \) of the \( N \)-point Gaussian quadrature? The solution is given by the following Golub-Welsch algorithm.

**Theorem 2** (Golub and Welsch 1969) For each \( n \geq 1 \), define \( \alpha_n, \beta_n \) by
\[ \alpha_n = \frac{\langle x_{p_{n-1}}, p_{n-1} \rangle}{\|p_{n-1}\|^2}, \quad \beta_n = \frac{\|p_n\|}{\|p_{n-1}\|} > 0. \]

Define the \( N \times N \) symmetric tridiagonal matrix \( T_N \) as in (2). Then the Gaussian quadrature nodes \( \{x_n\}_{n=1}^N \) are eigenvalues of \( T_N \). Letting \( v_n = (v_{n1}, \ldots, v_{nm})' \) be an eigenvector of \( T_N \) corresponding to eigenvalue \( x_n \), then the weights \( \{w_n\}_{n=1}^N \) are given by
\[
\begin{align*}
    w_n &= \frac{v_n^2}{\|v_n\|^2} \int w(x) \, dx > 0. \\
\end{align*}
\] (8)

**Proof** By (6) and the definition of \(\alpha_n, \beta_n\), for all \(n \geq 0\) we have
\[
p_{n+1}(x) = (x - \alpha_{n+1})p_n(x) - \beta_n^2 p_{n-1}(x).
\]

Note that this is true for \(n = 0\) by defining \(p_{-1}(x) = 0\) and \(\beta_0 = 0\). For each \(n\), let \(p_n^*(x) = p_n(x)/\|p_n\|\) be the normalized orthogonal polynomial. Then the above equation becomes
\[
\|p_{n+1}\|p_{n+1}^*(x) = \|p_n\|(x - \alpha_{n+1})p_n^*(x) - \|p_{n-1}\|\beta_n^2 p_{n-1}^*(x).
\]

Dividing both sides by \(\|p_n\| > 0\), using the definition of \(\beta_n, \beta_{n+1}\), and rearranging terms, we obtain
\[
\beta_n p_{n-1}^*(x) + \alpha_{n+1} p_n^*(x) + \beta_{n+1} p_{n+1}^*(x) = xp_n^*(x).
\]

In particular, setting \(x = x_k\) (where \(x_k\) is a root of \(p_N\)), we obtain
\[
\beta_n p_{n-1}^*(x_k) + \alpha_{n+1} p_n^*(x_k) + \beta_{n+1} p_{n+1}^*(x_k) = x_k p_n^*(x_k).
\]

for all \(n\) and \(k = 1, \ldots, N\). Since \(\beta_k = 0\) by definition and \(p_n^*(x_k) = 0\) (since \(x_k\) is a root of \(p_N\) and hence \(p_n^* = p_N/\|p_N\|\)), letting \(P(x) = (p_0^*(x), \ldots, p_{N-1}^*(x))'\) and collecting the above equation into a vector, we obtain
\[
T_N P(x_k) = x_k P(x_k)
\]

for \(k = 1, \ldots, N\). Define the \(N \times N\) matrix \(P\) by \(P = (P(x_1), \ldots, P(x_N))\). Then \(T_N P = \text{diag}(x_1, \ldots, x_N)P\), so \(x_1, \ldots, x_N\) are eigenvalues of \(T_N\) provided that \(P\) is invertible. Now since \(\{p_n^*(x)\}_{n=0}^{N-1}\) are normalized and Gaussian quadrature integrates all polynomials of degree up to \(2N - 1\) exactly, we have
\[
\delta_{mn} = (p_m^*, p_n^*) = \int w(x)p_m^*(x)p_n^*(x) \, dx = \sum_{k=1}^N w_k p_m^*(x_k)p_n^*(x_k)
\]

for \(m, n \leq N - 1\). Letting \(W = \text{diag}(w_1, \ldots, w_N)\), this equation becomes \(P W P' = I\). Therefore \(P, W\) are invertible and \(x_1, \ldots, x_N\) are eigenvalues of \(T_N\). Solving for \(W\) and taking the inverse, we obtain
\[
W^{-1} = P'P \iff \frac{1}{w_n} = \sum_{k=0}^{N-1} p_k^*(x_n)^2 > 0
\]

for all \(n\). To show (8), let \(v_n\) be an eigenvector of \(T_N\) corresponding to eigenvalue \(x_n\). Then \(v_n = cP(x_n)\) for some constant \(c \neq 0\). Taking the norm, we obtain
\[
\|v_n\|^2 = c^2 \|P(x_n)\|^2 = c^2 \sum_{k=0}^{N-1} p_k^*(x_n)^2 = \frac{c^2}{w_n} \iff w_n = \frac{c^2}{\|v_n\|^2}.
\]
Comparing the first element of $v_n = cP(x_n)$, noting that $p_0(x) = 1$ and hence $p_0^* = p_0/\|p_0\| = 1/\|p_0\|$, we obtain

$$c^2 = v_{n1}^2 \|p_0\|^2 = v_{n1}^2 \int w(x)p_0(x)^2 \, dx = v_{n1}^2 \int w(x) \, dx,$$

which implies (8). □

References

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. Journal of Finance, 59(4), 1481–1509. https://doi.org/10.1111/j.1540-6261.2004.00670.x.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. Quarterly Journal of Economics, 121(3), 823–866. https://doi.org/10.1162/qjec.121.3.823.

Curtiss, J. H. (1942). A note on the theory of moment generating functions. Annals of Mathematical Statistics, 13(4), 430–433. https://doi.org/10.1214/aoms/1177731541.

Farmer, L. E., & Toda, A. A. (2017). Discretizing nonlinear, non-Gaussian Markov processes with exact conditional moments. Quantitative Economics, 8(2), 651–683. https://doi.org/10.3982/QE737.

Flodén, M. (2008). A note on the accuracy of Markov-chain approximations to highly persistent AR(1) processes. Economics Letters, 99(3), 516–520. https://doi.org/10.1016/j.econlet.2007.09.040.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. Quarterly Journal of Economics, 127(2), 645–700. https://doi.org/10.1093/qje/qjs001.

Golub, G. H., & Welsch, J. H. (1969). Calculation of Gauss quadrature rules. Mathematics of Computation, 23(106), 221–230. https://doi.org/10.1090/S0025-5718-69-99647-1.

Lucas, Jr., R. E. (1978). Asset prices in an exchange economy. Econometrica, 46(6), 1429–1445. https://doi.org/10.2307/1913837.

Miller, III, Allen C., & Rice, T. R. (1983). Discrete approximations of probability distributions. Management Science, 29(3), 352–362. https://doi.org/10.1287/mnsc.29.3.352.

Pohl, W., Schmedders, K., & Wilms, O. (2018). Higher-order effects in asset pricing models with long-run risks. Journal of Finance, 73(3), 1061–1111. https://doi.org/10.1111/jofi.12615.

Rietz, T. A. (1988). The equity risk premium: A solution. Journal of Monetary Economics, 22(1), 117–131. https://doi.org/10.1016/0304-3932(88)90172-9.

Rouwenhorst, K. Geert. (1995). Asset pricing implications of equilibrium business cycle models. In F. C. Thomas (Ed.), Frontiers of business cycle research (pp. 294–330). Princeton: Princeton University Press.

Tanaka, K., & Toda, A. A. (2013). Discrete approximations of continuous distributions by maximum entropy. Economics Letters, 118(3), 445–450. https://doi.org/10.1016/j.econlet.2012.12.020.

Tanaka, K., & Toda, A. A. (2015). Discretizing distributions with exact moments: Error estimate and convergence analysis. SIAM Journal on Numerical Analysis, 53(5), 2158–2177. https://doi.org/10.1137/140971269.

Tauchen, G. (1986). Finite state Markov-chain approximations to univariate and vector autoregressions. Economics Letters, 20(2), 177–181. https://doi.org/10.1016/0165-1765(86)90168-0.

Tauchen, G., & Hussey, R. (1991). Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. Econometrica, 59(2), 371–396. https://doi.org/10.2307/2938261.

Toda, A. A., & Walsh, K. J. (2019). The equity premium and the one percent. Review of Financial Studies,. https://doi.org/10.1093/rfs/hhz121.

Tsionas, E. G. (2003). Exact solution of asset pricing models with arbitrary shock distributions. Journal of Economic Dynamics and Control, 27(5), 843–851. https://doi.org/10.1016/S0165-1889(02)00017-9.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.