UvA ECONOMETRICS

Discussion Paper: 2010/08

Saddlepoint approximation of expected shortfall for transformed means

Simon A. Broda and Marc S. Paolella

www.feb.uva.nl/ke/UvA-Econometrics

Amsterdam School of Economics
Department of Quantitative Economics
Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

Saddlepoint Approximation of Expected Shortfall for Transformed Means

Simon A. Broda[a,*] Marc S. Paolella[b,c]

[a]*Department of Quantitative Economics, University of Amsterdam, The Netherlands*

[b]*Swiss Banking Institute, University of Zurich, Switzerland*

[c]*Swiss Finance Institute*

December 2010

Abstract

Expected shortfall, as a coherent risk measure, has received a substantial amount of attention in the literature recently. For many distributions of practical interest however, it cannot be obtained in explicit form, and numerical techniques must be employed. The present manuscript derives a saddlepoint approximation for the expected shortfall associated with certain random variables that permit a stochastic representation in terms of some underlying random variables possessing a moment generating function. The new approximation can be evaluated quickly and reliably, and provides excellent accuracy. The doubly noncentral t distribution is considered as an example.

Key Words: Expected Shortfall, Noncentral t Distribution, Saddlepoint Approximation

*Address correspondence to Simon Broda, Department of Quantitative Economics, 1018WB Amsterdam, The Netherlands, or email s.a.broda@uva.nl. Part of the research of M. Paolella has been carried out within the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK), which is a research program supported by the *Swiss National Science Foundation*.

1 Introduction

The question of how to adequately measure risk is central to risk management, and has seen much controversy lately: beginning with the seminal paper of Artzner, Delbaen, Eber, and Heath (1999), an enormous body of literature has accumulated which criticizes the most widely used risk measure, Value at Risk (VaR), for its failure to be coherent. In particular, VaR is not subadditive, thus violating one of the four axioms defining a coherent risk measure. The expected shortfall, as a coherent alternative, appears to be emerging as a new standard.

Suppose the return on a financial position is represented by a continuous random variable X with density $f_X(x)$. Then the expected shortfall (ES) is defined, for a given confidence level $q \in (0, 1)$, as

$$\mathrm{ES}^{(\mathrm{q})}(\mathrm{X}) \equiv -\mathbb{E}\left[X | X \leq x_q\right] = -\frac{1}{q} \int_{-\infty}^{x_q} x f_X(x) \mathrm{d}x, \tag{1}$$

where x_q is the corresponding VaR, i.e., the $100q\%$ quantile of X, see Acerbi and Tasche (2002a,b). Often q is set to 1%. Note that here, we have defined the expected shortfall in terms of relative loss, as is common. For many distributions of interest, (1) must be evaluated numerically. This time-consuming task potentially precludes the use of certain distributions in computationally demanding tasks such as mean-ES portfolio allocation, so that a reliable approximation is desirable. The saddlepoint approximation, introduced next, is a natural candidate.

Saddlepoint approximations for densities result from expanding the Fourier-Mellin integral of the moment generating function (hereafter mgf) in an asymptotic series. Since their introduction by Daniels (1954), they have been found to deliver excellent accuracy in countless applications; see Butler (2007), Broda and Paolella (2011), and the references therein. The cumulative distribution function could be approximated by numerically integrating the approximate density; however, Lugannani and Rice (1980) derived a direct approximation to the distribution function which circumvents this need. Both Daniels's and Lugannani and Rice's approximation require that the random variable to which the approximation is to be applied possess an mgf. Daniels and Young (1991) have generalized these results to marginal distributions of nonlinear functions of bivariate random vectors; see also Jing and Robinson (1994).

Concerning the expected shortfall as defined in (1) above, the result of Martin (2006) can be used whenever the moment generating function of X exists. For certain problems however, X does not possess an mgf, but may permit a stochastic representation in terms of random variables that do. The present paper derives a saddlepoint approximation that is applicable in this setting. At first glance, it might perhaps appear as if the most natural route to such a result were via a routine application of asymptotic expansions. This is certainly possible (if a bit tedious), and would lead to a result which generalizes that of Martin (2006) in much the same the way Daniels and Young's result generalizes that of Lugannani and Rice. However, careful scrutiny of the results available in the univariate setting reveals that an alternative strategy may lead to an approximation that is both more accurate and simpler to compute. This is the approach taken here.

The plan is as follows. Section 2 reviews some useful results from the theory of saddlepoint approximations. Section 3 demonstrates an alternative method of proof for some known saddlepoint approximations to expected shortfall, and motivates the new approximation given in Section 4. Section 5 applies the approximation to the noncentral t distribution and concludes.

2 Saddlepoint Approximations for Densities and Tail Probabilities

Consider the mean of n independent and identically distributed variables, $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, and suppose that each X_i possesses a density, and that its cumulant generating function $\mathbb{K}_X(t)$ converges in a nonvanishing interval containing the origin. The cumulant generating function is the natural logarithm of the mgf, which is, in turn, defined as

$$\mathbb{M}_X(t) \equiv \mathbb{E}\left[e^{tX}\right]$$

whenever the expectation exists. It generates the moments of a distribution, in the sense that its jth derivative evaluated at zero equals the jth moment of X. The usefulness of the mgf in the context of risk management derives from the fact that the mgf corresponding to the convolution of two distributions (viz., the mgf of a sum of two independent random variables) is given by the product of the individual mgfs. Hence, the mgf possesses a tractable form in many cases in which the distribution function does not. If the distribution of $\bar{X}$ is absolutely continuous, as we shall assume throughout, then its density $f_n(\bar{x})$ can be recovered from the mgf via the Fourier-Mellin integral

$$f_n(\bar{x}) = \frac{n}{2\pi\mathrm{i}}\int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} e^{n(\mathbb{K}_X(t)-tx)}\mathrm{d}t. \tag{2}$$

Similar expressions exist for the distribution function and the expected shortfall. Typically, the integral in (2) permits no analytical solution, and numerical quadrature must be used. An alternative is to approximate the integral by means of a truncated asymptotic series: Daniels (1954) derived the asymptotic expansion

$$f_n(\bar{x}) = g_n(\bar{x})\left[\sum_{k=0}^{m}\frac{a_k}{n^k} + \mathcal{O}\left(n^{-(m+1)}\right)\right], \tag{3}$$

where

$$g_n(\bar{x}) = \sqrt{\frac{n}{2\pi\,\mathbb{K}_X''(\hat{t})}}e^{-n\hat{w}^2/2}, \quad \hat{w} = \operatorname{sgn}\left(\hat{t}\right)\sqrt{2\hat{t}\bar{x} - 2\mathbb{K}_X(\hat{t})},$$

and the saddlepoint $\hat{t} = \hat{t}\left(\bar{x}\right)$ solves

$$\bar{x} = \mathbb{K}_X'(\hat{t})\,.$$

In most applications this has to be evaluated numerically for each $\bar{x}$ at which the density is to be approximated. The first two coefficients in the expansion (3) are

$$a_0 = 1 \quad \text{and } a_1 = \frac{1}{8}\hat{\lambda}_4 - \frac{5}{24}\hat{\lambda}_3^2,$$

where $\hat{\lambda}_j = \mathbb{K}_X^{(j)} / (\mathbb{K}_X'')^{j/2}$, and the truncated series g_n is referred to as the saddlepoint approximation to f_n.

A uniform asymptotic expansion for the distribution function of $\bar{X}$ has been derived in Lugannani and Rice (1980). It is given by

$$F_n(\bar{x}) = \Phi\left(\hat{w} n^{1/2}\right) + \sqrt{\frac{1}{2\pi n}} e^{-n\hat{w}^2/2} \left[\sum_{k=0}^{m} \frac{b_k}{n^k} + \mathcal{O}\left(n^{-(m+1)}\right)\right], \tag{4}$$

where Φ is the cdf of the standard normal distribution. The first two coefficients in the expansion (4) are

$$b_0 = \frac{1}{\hat{w}} - \frac{1}{\hat{u}} \quad \text{and } b_1 = \frac{1}{\hat{u}^3} - \frac{1}{\hat{w}^3} + \frac{\hat{\lambda}_3}{2\hat{u}^2} - \frac{a_1}{\hat{u}},$$

where

$$\hat{u} = \hat{t}\sqrt{\mathbb{K}_X''(\hat{t})}$$

and a_1 is as in (3).

Lugannani and Rice derived expansion (4) using complex variable methods. It can, however, also be obtained from (3) by means of the following result of Temme (1982), which we shall use frequently.

Theorem 1 (Temme, 1982). *Suppose $\Psi_n(\zeta)$ is analytic in a strip containing the real axis in its interior and such that, for some fixed λ_k and ω,*

$$\frac{\mathrm{d}^k}{\mathrm{d}\zeta^k}\Psi_n(\zeta) = \mathcal{O}\left(\left|\zeta^{\lambda_k}\right| e^{\omega\zeta^2}\right) \quad \text{as } \operatorname{Re}\zeta \to \infty.$$

If $\Psi_n(\zeta)$ permits the asymptotic expansion

$$\Psi_n(\zeta) = \sum_{k=0}^{m} \frac{\psi_k(\zeta)}{n^k} + \mathcal{O}\left(\Psi_n(\zeta) n^{-(m+1)}\right),$$

where the $\psi_k(\zeta)$ are analytic in ζ and do not depend on n, then

$$\begin{aligned} H_n(\eta) = \sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\eta} e^{-n\zeta^2/2} \Psi_n(\zeta) \mathrm{d}\zeta = \\ = \Phi\left(\eta n^{1/2}\right)\left[\sum_{k=0}^{m} \frac{A_k}{n^k} + \mathcal{O}(n^{-(m+1)})\right] + \sqrt{\frac{1}{2\pi n}} e^{-n\eta^2/2} \left[\sum_{k=0}^{m} \frac{B_k(\eta)}{n^k} + \mathcal{O}\left(n^{-(m+1)}\right)\right], \end{aligned}$$

where A_k are the coefficients in the asymptotic expansion

$$H_n(\infty) = \sum_{k=0}^{\infty} \frac{A_k}{n^k},$$

$$B_{-1}(\eta) = 0, \quad \textit{and } \eta B_k(\eta) = A_k - \psi_k(\eta) + \frac{\mathrm{d}}{\mathrm{d}\eta} B_{k-1}(\eta), \quad k \in \{0, 1, \ldots\}.$$

3 Saddlepoint Approximations for Expected Shortfall

The first step to approximating the expected shortfall of $\bar{X}$ is to find an approximation to its quantiles $\bar{x}_q$. This can be achieved by truncating the series for $F_n(\bar{x})$ after k terms, yielding the approximation $\hat{F}_n^k(\bar{x})$, say, and numerically solving

$$\hat{F}_n^k(\hat{\bar{x}}_q^k) = q, \quad q \in (0,1). \tag{5}$$

It then remains to find an approximation to integrals of the form

$$I_n(c) = \int_{-\infty}^{c} \bar{x} f_n(\bar{x}) \mathrm{d}\bar{x}. \tag{6}$$

Martin (2006) shows that

$$I_n(c) = \Phi(\hat{w}_c n^{1/2})\mu + \sqrt{\frac{1}{2\pi n}} e^{-n\hat{w}_c^2/2} \left(\frac{\mu}{\hat{w}_c} - \frac{c}{\hat{u}_c} + \mathcal{O}\left(n^{-1}\right) \right), \tag{7}$$

where $\mu = \mathbb{K}_X'(0)$, and hatted quantities with a subscript c correspond to those in (3) and (4), but evaluated at c rather than $\bar{x}$. Martin derived his approximation using the same technique employed by Lugannani and Rice in obtaining their expansion of the distribution function. The same result is obtained by applying Theorem 1 to the integral in (6) and retaining only the leading term. Incorporating the second term, which is derived in Appendix A, one has

$$\begin{aligned} I_n(c) = {} & \Phi(\hat{w}_c n^{1/2})\mu + \\ & + \sqrt{\frac{1}{2\pi n}} e^{-n\hat{w}_c^2/2} \left(\frac{\mu}{\hat{w}_c} - \frac{c}{\hat{u}_c} + \frac{1}{n}\left[\frac{c}{\hat{u}_c^3} - \frac{\mu}{\hat{w}_c^3} + \frac{c\hat{\lambda}_{3,c}}{2\hat{u}_c^2} - \frac{c a_1}{\hat{u}_c} - \frac{1}{\hat{t}_c \hat{u}_c} \right] + \mathcal{O}\left(n^{-2}\right) \right), \end{aligned} \tag{8}$$

where a_1 is as in (3).

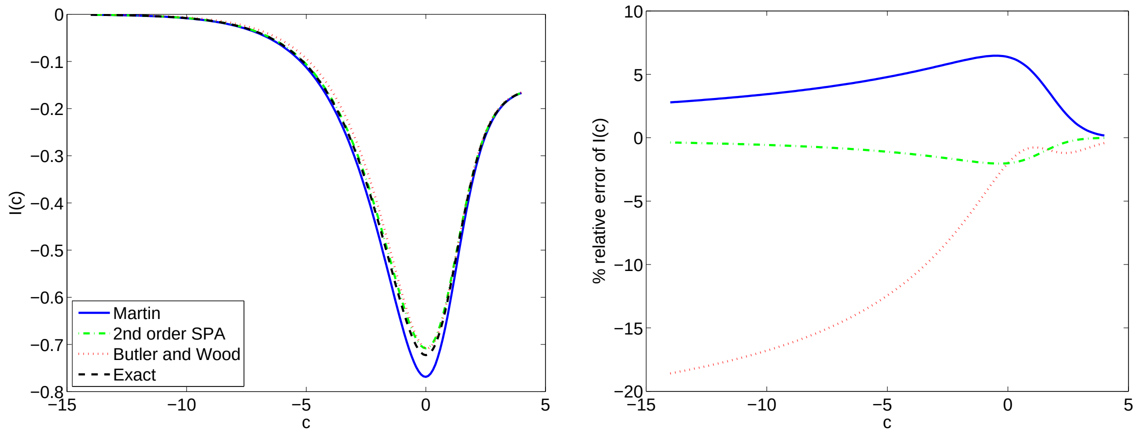

Figure 1: Approximations to $I_n(c)$ (left panel) and relative error in percent (right panel).

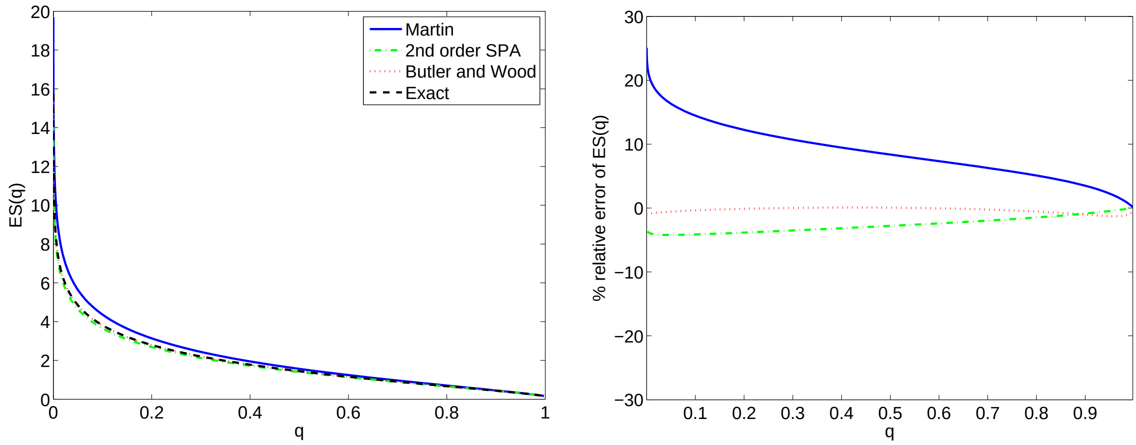

Figure 2: Approximations (left panel) and relative error in percent (right panel) for the expected shortfall at level q.

Butler and Wood (2004), via a completely different method of proof, obtained an approximation for the mgf, and its logarithmic derivatives, of the truncated random variable ${}_a\bar{X}_b$ with density $f_n(\bar{x})\mathbb{I}_{(a,b)}(\bar{x})/(F_n(b)-F_n(a))$, where as before, f_n and F_n are the density and distribution function of $\bar{X}$, respectively. Setting $a=-\infty$, $b=c$, and evaluating their approximation for the logarithmic derivative at zero produces the following approximation:

$$I_n(c) = \Phi(\hat{w}_c n^{1/2})\mu + \sqrt{\frac{1}{2\pi n}} e^{-n\hat{w}_c^2/2} \left(\frac{\mu}{\hat{w}_c} - \frac{c}{\hat{u}_c} + \frac{1}{n}\left[\frac{c}{\hat{w}_c^3} - \frac{\mu}{\hat{w}_c^3} - \frac{1}{\hat{t}\hat{u}_c} \right] + \mathcal{O}\left(n^{-1}\right) \right). \quad (9)$$

We illustrate the relative merits of approximations (7), (8) and (9) by considering one copy (i.e., $n=1$) of a Normal Inverse Gaussian (NIG) random variable with parameters $(\alpha,\beta,\delta,\mu) = (1,-1/2,2,0)$. The NIG belongs to the family of Generalized Hyperbolic (GHyp) distributions. These, and the NIG in particular, have been very successfully applied to financial returns data, see, e.g., Eberlein and Keller (1995), Barndorff-Nielsen (1997) and Broda and Paolella (2009). As shown in the latter paper, the saddlepoint approximation for the NIG is explicit, and the relevant quantities are $\hat{t}_c = z\alpha/\bar{y} - \beta$, $z = (c-\mu)/\delta$, $\bar{y} = \sqrt{1+z^2}$, $\mathbb{K}_X''(\hat{t}_c) = \bar{y}^3\delta^3\alpha$, $\hat{w}_c = \operatorname{sgn}(\hat{t}_c)[2\delta(\bar{y}\alpha - z\beta - (\alpha^2-\beta^2)^{1/2}]^{1/2}$, $\hat{\lambda}_{3,c} = 3z(\bar{y}\alpha\delta)^{-1/2}$ and $\hat{\lambda}_{4,c} = 3(1+5z^2)(\bar{y}\alpha\delta)^{-1}$.

The results are shown in Figure 1. As expected, the second order approximation (8) clearly dominates the first order approximation (7), with under 10% error across the entire support. It is also observed that despite being formally accurate to the same order, approximation (9) cannot compete even with (7). This picture changes dramatically if instead of the tail integral $I_n(c)$, one considers its normalized version, viz., the expected shortfall: the approximation is

$$\mathrm{ES}^{(\mathrm{q})}(\mathrm{X}) \sim -\frac{I_n^{\ell}(\hat{\hat{x}}_q^k)}{\hat{F}^k(\hat{\hat{x}}_q^k)},$$

where $\ell=1$ refers to approximation (7), $\ell=2$ to approximation (8), and $\ell=3$ to approximation (9), and $\hat{\hat{x}}_q^k$ is determined according to (5) with $k=1$ if $\ell\in\{1,3\}$ and $k=2$ if $\ell=2$. In other words, the order of the approximation to the distribution function used in determining the quantile is chosen to agree with that of the approximation to I_n.

The approximations are illustrated in Figure 2. Despite the fact that (9) is only an expansion to order $n^{-3/2}$, the resulting approximation to the expected shortfall is comparable to that based on (8), which is $\mathcal{O}\left(n^{-5/2}\right)$. The reason for this is the following: In order to benefit from error cancellation, one would ideally approximate the expected shortfall by

$$\mathrm{ES}^{(\mathrm{q})}(\mathrm{X}) \sim -\int_{-\infty}^{\hat{\bar{x}}_q^k} \bar{x} \hat{f}_n^{\ell}(\bar{x}) \mathrm{d}\bar{x} \left[\int_{-\infty}^{\hat{\bar{x}}_q^k} \hat{f}_n^{\ell}(\bar{x}) \mathrm{d}\bar{x}\right]^{-1},$$

that is, use the same approximation $\hat{f}_n^{\ell}(\bar{x})$ in both the numerator and denominator. If the first order approximation to the distribution function, $\hat{F}^1(\hat{\bar{x}}_q^k)$, is used as the denominator, this amounts to approximating the numerator integral by applying Theorem 1 to

$$-\int_{-\infty}^{\hat{\bar{x}}_q^k} \bar{x} \left[\frac{\mathrm{d}}{\mathrm{d}\bar{x}} \hat{F}_n^1(\bar{x})\right] \mathrm{d}\bar{x}.$$

Now, differentiating the distribution function approximation yields

$$\frac{\mathrm{d}}{\mathrm{d}\bar{x}} \hat{F}_n^1(\bar{x}) = g_n(\bar{x}) \left(1 + \frac{c_1}{n}\right),$$

where

$$c_1 = \frac{1}{\hat{u}^2} + \frac{\hat{\lambda}_3}{2\hat{u}} - \frac{\hat{u}}{\hat{w}^3}.$$

A simple way of seeing that this approach leads to approximation (9) is to replace a_1 with $c_1 + \mathcal{O}\left(n^{-1}\right)$ in the second-order expansion (8).

4 Expected Shortfall for Transformed Means

The approximations given in the foregoing section are limited to random variables which possess an mgf. This section derives an approximation for the case where the random variable of interest does not possess an mgf, but can be represented as a smooth function of two underlying random variables that do. The derivation is motivated by the following observation: as demonstrated above for the NIG distribution, the expected shortfall approximation based on (9) provides similar accuracy to that based on (8), despite the fact that the former only involves quantities which are anyway needed for the saddlepoint approximation to the distribution function (4). Note that prior to computing the expected shortfall, one needs to find the relevant quantile of the loss distribution, and thus evaluate the approximate distribution function. Therefore computing the expected shortfall via approximation (9), unlike using (8), entails no extra effort.

To fix ideas, let $X = (X_1, X_2)$ be a bivariate random vector having a density and a joint cumulant generating function, and denote the latter as $\mathbb{K}_X(\mathbf{t})$. Denote by $\mathbb{K}_X'(\mathbf{t})$ and $\mathbb{K}_X''(\mathbf{t})$ its gradient and hessian, respectively, and by $\bar{X} = (\bar{X}_1, \bar{X}_2)$ the mean of X from a sample of n

observations. Consider a smooth bijection g mapping the support of Y onto the support of X, such that $\bar{X} = g(Y) = (g_1(Y), g_2(Y))'$, and denote its inverse by h so that $Y = (Y_1, Y_2) = h(\bar{X}) = (h_1(\bar{X}), h_2(\bar{X}))'$. We wish to obtain an approximation for the expected shortfall of Y_1 which is analogous to (9); that is, we will approximate

$$I_n(c) = \int_{-\infty}^{c} \bar{x} f_n(y_1) \mathrm{d}\bar{x} \tag{10}$$

by replacing $f_n(y_1)$ with the first derivative of the appropriate distribution approximation and then applying Theorem 1.

Our starting point is the following result from Daniels and Young (1991). Let $\nabla_{y_i} g(y) = (\partial g_1/\partial y_i, \partial g_2/\partial y_i)', i \in \{1, 2\}$, and let $J_g(y) = \big(\partial g/\partial \mathbf{y}\big)$ denote the Jacobian of g. The density and distribution function of Y_1 are

$$f_n(y_1) = \sqrt{\frac{n}{2\pi\tilde{K}}} e^{-n\tilde{w}^2/2} \left[1 + \mathcal{O}\left(n^{-1}\right)\right]$$

and

$$F_n(y_1) = \Phi\left(\tilde{w}n^{1/2}\right) + \sqrt{\frac{1}{2\pi n}} e^{-n\tilde{w}^2/2} \left[\frac{1}{\tilde{w}} - \frac{1}{\tilde{u}} + \mathcal{O}\left(n^{-1}\right)\right], \tag{11}$$

respectively, where $\tilde{w} = \mathrm{sgn}(y_1 - \alpha)\sqrt{2\big(\tilde{\mathbf{t}}' g(\tilde{\mathbf{y}}) - \mathbb{K}_X(\tilde{\mathbf{t}})\big)}$, $\tilde{\mathbf{y}} = (y_1, \tilde{y}_2)$, $\alpha = h_1(\boldsymbol{\mu})$, $\boldsymbol{\mu} = \mathbb{K}_X'(\mathbf{0})$,

$$\tilde{K} = \frac{\det\left(\mathbb{K}_X''(\tilde{\mathbf{t}})\right)}{\det^2 J_g(\tilde{\mathbf{y}})} \left[\nabla_{y_2} g(\tilde{y})' \left(\mathbb{K}_X''(\tilde{\mathbf{t}})\right)^{-1} \nabla_{y_2} g(\tilde{\mathbf{y}}) + \tilde{\mathbf{t}}' \nabla_{y_2}^2 g(\tilde{\mathbf{y}})\right],$$

$\tilde{u} = \big(\tilde{\mathbf{t}}' \nabla_{y_1} g(\tilde{\mathbf{y}})\big)\tilde{K}^{1/2}$, and, for each value of y_1, $\tilde{\mathbf{t}}$ and $\tilde{y}_2$ solve the system

$$\begin{aligned} \mathbb{K}_X'(\tilde{\mathbf{t}}) &= g(\tilde{\mathbf{y}}) \\ \tilde{\mathbf{t}}' \nabla_{y_2} g(\tilde{\mathbf{y}}) &= 0. \end{aligned}$$

The next step is to differentiate the expansion for F_n. This yields

$$f_n(y_1) = \sqrt{\frac{n}{2\pi\tilde{K}}} e^{-n\tilde{w}^2/2} \left[1 - \frac{1}{n}\left[\frac{\tilde{u}}{\tilde{w}^3} + \tilde{K}^{1/2} \frac{\mathrm{d}}{\mathrm{d}y_1}\frac{1}{\tilde{u}}\right] + \mathcal{O}\left(n^{-1}\right)\right].$$

As shown in Appendix B, substituting this into (10) and applying Theorem 1 yields

$$\begin{aligned} I_n(c) = {} & \Phi(\tilde{w}_c n^{1/2}) \left(A_0 + \frac{1}{n} A_1\right) + \\ & + \sqrt{\frac{1}{2\pi n}} e^{-n\tilde{w}_c^2/2} \left(\frac{A_0}{\tilde{w}_c} - \frac{c}{\tilde{u}_c} + \frac{1}{n}\left[\frac{A_1}{\tilde{w}_c} + \frac{c}{\tilde{w}_c^3} - \frac{A_0}{\tilde{w}_c^3} - \frac{1}{\big(\tilde{\mathbf{t}}_c' \nabla_{y_1} g(\tilde{\mathbf{y}}_c)\big)\tilde{u}_c}\right] + \mathcal{O}\left(n^{-1}\right)\right), \end{aligned} \tag{12}$$

where as before, tilded quantities with a subscript c are evaluated at c rather than $\bar{x}$, and A_0 and A_1 are coefficients in the asymptotic expansion of $\mathbb{E}[y_1]$. Denoting the elements of the Hessian of $h_1(\bar{X})$ evaluated at $\boldsymbol{\mu}$ as h^{ij} and the elements of $\mathbb{K}_X''(\mathbf{0})$ as κ_{ij}, these are $A_0 = \alpha$

and $A_1 = h^{ij}\kappa_{ij}/2$, see Hurt (1976). Here we have used the summation convention, i.e., indices appearing as both a subscript and a superscript are to be summed over. For practical purposes, saddlepoint approximations are often applied for $n = 1$, incorporating the convolution into the mgf itself. Following that practice, we have the following result.

THEOREM 2. *Let $X = (X_1, X_2)$ be a bivariate random vector possessing a density and joint cumulant generating function $\mathbb{K}_X(\mathbf{t})$. Let g be a smooth bijection such that $X = g(Y) = (g_1(Y), g_2(Y))'$, with inverse $Y = (Y_1, Y_2) = h(X) = (h_1(X), h_2(X))'$. Then*

$$\mathbb{E}\left[Y_1 \middle| Y_1 < c\right] \approx (\alpha + A_1) + \frac{\phi(\tilde{w}_c)}{\hat{F}^1(c)} \left((c-\alpha)\left(\frac{1}{\tilde{w}_c^3} - \frac{1}{\tilde{u}_c}\right) - \frac{1}{\left(\tilde{\mathbf{t}_c}'\nabla_{y_1} g(\tilde{\mathbf{y}}_c)\right)\tilde{u}_c} + \frac{A_1}{\tilde{u}_c} \right), \tag{13}$$

where $\tilde{w}_c = \operatorname{sgn}(c-\alpha)\sqrt{2\left(\tilde{\mathbf{t}_c}'g(\tilde{\mathbf{y}}_c) - \mathbb{K}_X(\tilde{\mathbf{t}_c})\right)}$, $\tilde{\mathbf{y}}_c = (c, \tilde{y}_{2,c})$, $\alpha = h_1(\boldsymbol{\mu})$, $\boldsymbol{\mu} = \mathbb{K}_X'(\mathbf{0})$, $A_1 = h^{ij}\kappa_{ij}/2$, h^{ij} and κ_{ij} are the elements of the Hessian of $h_1(\bar{X})$ at $\boldsymbol{\mu}$ and $\mathbb{K}_X''(\mathbf{0})$, respectively,

$$\tilde{K}_c = \frac{\det\left(\mathbb{K}_X''(\tilde{\mathbf{t}}_c)\right)}{\det^2 J(\tilde{\mathbf{y}}_c)} \left[\nabla_{y_2} g(\tilde{y}_c)' \left(\mathbb{K}_X''(\tilde{\mathbf{t}}_c)\right)^{-1} \nabla_{y_2} g(\tilde{\mathbf{y}}_c) + \tilde{\mathbf{t}}_c' \nabla_{y_2}^2 g(\tilde{\mathbf{y}}_c)\right], \quad \tilde{u}_c = \left(\tilde{\mathbf{t}}_c'\nabla_{y_1} g(\tilde{\mathbf{y}}_c)\right)\tilde{K}_c^{1/2},$$

$\hat{F}^1(c) = \Phi(\tilde{w}_c) + \phi(\tilde{w}_c)\left[\tilde{w}_c^{-1} - \tilde{u}_c^{-1}\right]$, and $\tilde{\mathbf{t}}_c$ and $\tilde{y}_{2,c}$ solve the system

$$\begin{aligned} \mathbb{K}_X'(\tilde{\mathbf{t}}_c) &= g(\tilde{\mathbf{y}}_c) \\ \tilde{\mathbf{t}}_c'\nabla_{y_2} g(\tilde{\mathbf{y}}_c) &= 0. \end{aligned}$$

5 Application

In this section, we apply our approximation to the doubly noncentral t distribution, which has recently found application as a model for financial returns, see Harvey and Siddique (1999), Broda and Paolella (2007), Paolella (2010), and the references therein. Despite the excellent fit it offers, the widespread adoption of this distribution has been hampered by the computational difficulties associated with it: neither the density nor the distribution function can be expressed in closed form. This limitation is overcome by the saddlepoint approximations for the density and distribution functions derived in Broda and Paolella (2007). The resulting approximations are explicit and hence trivially computed, while offering extraordinary accuracy. This benefit is inherited by the expected shortfall calculation as well.

Apart from location and scale parameters (which, without loss of generality, can be taken to be zero and one, respectively), the doubly noncentral t distribution has 3 parameters: the numerator noncentrality μ, denominator noncentrality θ, and the degrees of freedom, n. The relevant quantities entering the approximations are

$$\left(\tilde{\mathbf{t}}_c'\nabla_{y_1} g(\tilde{\mathbf{y}}_c)\right) = (\tilde{t}_{1,c}\tilde{y}_{2,c}), \qquad \tilde{K}_c = \left[(c^2 + 2n\tilde{t}_{2,c})(2n\nu^2 + 4\theta\nu^3) + 4n^2\tilde{y}_{2,c}^2\right] \Big/ \left(2n\tilde{y}_{2,c}^2\right)^2,$$

$$\tilde{w}_c = \sqrt{-\mu\tilde{t}_{1,c} - n\log\nu - 2\theta\nu\tilde{t}_{2,c}}\,\operatorname{sgn}(c-\alpha), \qquad \alpha = \mu/\sqrt{1+\theta/n},$$

where $\nu = (1-2\tilde{t}_{2,c})^{-1}$, $\tilde{t}_{1,c} = -\mu + c\tilde{y}_{2,c}$, $\tilde{t}_{2,c} = -(c\tilde{t}_{1c})/(2n\tilde{y}_{2,c})$, and, with $a_3 = c^4 + 2nc^2 + n^2$, $a_2 = -2c^3\mu - 2cn\mu$, $a_1 = c^2\mu^2 - nc^2 - n^2 - \theta n$, $a_0 = cn\mu$,

$$c_2 = \frac{a_2}{a_3}, \quad c_1 = \frac{a_1}{a_3}, \quad c_0 = \frac{a_0}{a_3}, \quad q = \frac{1}{3}c_1 - \frac{1}{9}c_2^2,$$

$$r = \frac{1}{6}(c_1c_2 - 3c_0) - \frac{1}{27}c_2^3, \quad m = q^3 + r^2, \quad \text{and} \quad s_{1,2} = (r \pm \sqrt{m})^{1/3},$$

$$\tilde{y}_{2,c} = \sqrt{-4q}\cos\left(\cos^{-1}\left(r/\sqrt{-q^3}\right)/3\right) - \frac{c_2}{3},$$

and it is easily seen that $A_1 = 3\alpha(n+2\theta)/(4(n+\theta)^2)$. With these, approximation (13) can be computed. For illustration, we choose the set of parameters $(\mu, \theta) = (-10, 10)$, which corresponds to a heavily left-skewed distribution. For $n \in \{2, 4, 8, 16\}$, Figure 3 displays the accuracy of (13). The exact values (the solid line in the graph) are obtained from the relationship

$$I_n(c) = cF_n(c) - \int_{-\infty}^{c} F_n(\bar{x})\mathrm{d}\bar{x}, \quad c < 0, \tag{14}$$

which follows from (6) upon integration by parts, and where the exact distribution function F_n is computed via the algorithm described in Reeve (1986). The accuracy of the approximation clearly improves with larger n, as expected, but is already quite good for n as low as 2 (note that for $n = 1$, the doubly noncentral t does not possess a mean, and thus the expected shortfall is not defined). When the t distribution is fitted to actual financial returns data, the degrees of freedom parameter is typically estimated between 4 and 8. The approximation is certainly accurate enough in this range for all practical purposes.

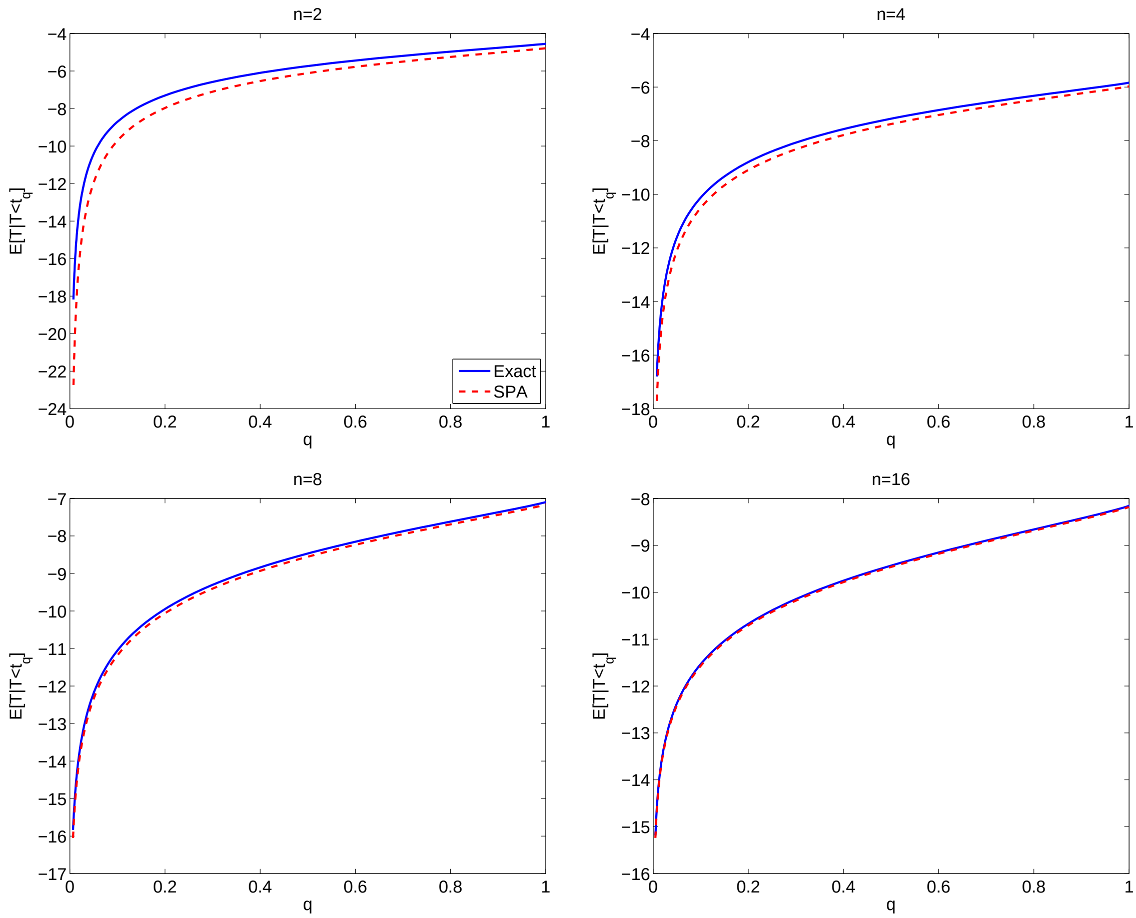

Figure 3: Approximations to expected shortfall of a doubly noncentral t with $\mu = -10$, $\theta = 10$, and $n = 2, 4, 8, 16$, from top left to bottom right.

References

Acerbi, C. and D. Tasche (2002a): "Expected Shortfall: A Natural Coherent Alternative to Value at Risk," *Economic Notes*, 31, 379–388. 1

——— (2002b): "On the coherence of Expected Shortfall," *Journal of Banking and Finance*, 26, 1487–1503. 1

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999): "Coherent Measures of Risk," *Mathematical Finance*, 9, 203–228. 1

Barndorff-Nielsen, O. E. (1997): "Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling," *Scandinavian Journal of Statistics*, 24, 1–13. 5

Broda, S. and M. S. Paolella (2007): "Saddlepoint Approximations for the Doubly Noncentral t Distribution," *Computational Statistics & Data Analysis*, 51, 2907–2918. 8

Broda, S. A. and M. S. Paolella (2009): "CHICAGO: A Fast and Accurate Method for Portfolio Risk Calculation," *Journal of Financial Econometrics*, 7, 412–436. 5

——— (2011): "Saddlepoint Approximations: A Review and Some New Applications," in *Handbook of Computational Statistics*, ed. by J. E. Gentle, W. K. Härdle, and Y. Mori, Berlin: Springer, 2nd ed. 1

Butler, R. W. (2007): *Saddlepoint Approximations with Applications*, Cambridge: Cambridge University Press. 1

Butler, R. W. and A. T. A. Wood (2004): "Saddlepoint Approximation for Moment Generating Functions of Truncated Random Variables," *Annals of Statistics*, 32, 2712–2730. 4

Daniels, H. E. (1954): "Saddlepoint Approximation in Statistics," *Ann. Math. Statist.*, 25, 631–650. 1, 2

Daniels, H. E. and G. A. Young (1991): "Saddlepoint Approximation for the Studentized Mean, with an Application to the Bootstrap," *Biometrika*, 78, 169–179. 1, 7

Eberlein, E. and U. Keller (1995): "Hyperbolic Distributions in Finance," *Bernoulli*, 1, 281–299. 5

Harvey, C. R. and A. Siddique (1999): "Autoregressive Conditional Skewness," *Journal of Financial and Quantitative Analysis*, 34, 465–487. 8

Hurt, J. (1976): "Asymptotic Expansions of Functions of Statistics," *Applications of Mathematics*, 21, 444–456. 8

Jing, B. and J. Robinson (1994): "Saddlepoint Approximations for Marginal and Conditional Probabilities of Transformed Variables," *Annals of Statistics*, 22, 1115–1132. 1

Lugannani, R. and S. O. Rice (1980): "Saddlepoint Approximations for the Distribution of Sums of Independent Random Variables," *Adv. Appl. Prob.*, 12, 475–490. 1, 3, 4

Martin, R. (2006): "The Saddlepoint Method and Portfolio Optionalities," *Risk Magazine*, 19, 93–95. 1, 4

Paolella, M. S. (2010): "ALRIGHT: Asymmetric Large-Scale (I)GARCH with Hetero-Tails," Tech. rep., available at SSRN. 8

Reeve, C. P. (1986): "An Algorithm for Computing the Doubly Noncentral t C.D.F. to a Specified Accuracy," SED Note N-86-5, National Bureau of Standards, Statistical Engineering Division. 9

Temme, N. M. (1982): "The Uniform Asymptotic Expansion of a Class of Integrals Related to Cumulative Distribution Functions," *SIAM Journal on Mathematical Analysis*, 13, 239–253. 3

A Proof of (8)

Replacing $f_n(\bar{x})$ in (6) with its second order saddlepoint approximation and applying Theorem 1 yields

$$\begin{aligned}
I_n(c) &= \int_{-\infty}^{c} \bar{x} f_n(\bar{x}) \mathrm{d}\bar{x} \\
&= \sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\hat{w}_c} \frac{\bar{x}}{\sqrt{\mathbb{K}_x''(\hat{t})}} e^{-n\hat{w}^2/2} \left[1 + \frac{a_1}{n} + \mathcal{O}\left(n^{-2}\right)\right] \mathrm{d}\bar{x} \\
&= \sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\hat{w}_c} \frac{\mathrm{d}\bar{x}}{\mathrm{d}\hat{w}} \frac{\bar{x}}{\sqrt{\mathbb{K}_x''(\hat{t})}} e^{-n\hat{w}^2/2} \left[1 + \frac{a_1}{n} + \mathcal{O}\left(n^{-2}\right)\right] \mathrm{d}\hat{w} \\
&= \sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\hat{w}_c} e^{-n\hat{w}^2/2} \Psi_n(\hat{w}) \mathrm{d}\hat{w},
\end{aligned}$$

where

$$\begin{aligned}
\Psi_n(\hat{w}) &\equiv \frac{\mathrm{d}\bar{x}}{\mathrm{d}\hat{w}} \frac{\bar{x}}{\sqrt{\mathbb{K}_x''(\hat{t})}} \left[1 + \frac{a_1}{n} + \mathcal{O}\left(n^{-2}\right)\right] \\
&= \frac{\hat{w}\bar{x}}{\hat{u}} \left[1 + \frac{a_1}{n} + \mathcal{O}\left(n^{-2}\right)\right] \\
&= \psi_0(\hat{w}) + \frac{\psi_1(\hat{w})}{n} + \mathcal{O}\left(n^{-2}\right),
\end{aligned} \tag{15}$$

with

$$\psi_0(\hat{w}) \equiv \frac{\hat{w}\bar{x}}{\hat{u}} \quad \text{and} \quad \psi_1(\hat{w}) \equiv \frac{\hat{w}\bar{x}}{\hat{u}} \left[\frac{1}{8}\hat{\lambda}_4 - \frac{5}{24}\hat{\lambda}_3^2\right].$$

Equation (15) is in the form of Theorem 1. The coefficients in the expansion evaluate to

$$B_0(\hat{w}_c) = \frac{\mu}{\hat{w}_c} - \frac{c}{\hat{u}_c}$$

and

$$\begin{aligned}
\hat{w}_c B_1(\hat{w}_c) &= A_1 - \psi_1(\hat{w}_c) + \frac{\mathrm{d}}{\mathrm{d}\hat{w}_c} B_0(\hat{w}_c) \\
&= -\frac{\hat{w}_c c}{\hat{u}_c}\left[\frac{1}{8}\hat{\lambda}_4 - \frac{5}{24}\hat{\lambda}_3^2\right] - \frac{\mu}{\hat{w}_c^2} - \frac{1}{\hat{u}_c}\frac{\mathrm{d}c}{\mathrm{d}\hat{w}_c} + \frac{c}{\hat{u}_c^2}\frac{\mathrm{d}}{\mathrm{d}\hat{w}_c}\hat{u}_c \\
&= -\frac{\hat{w}_c c}{\hat{u}_c}\left[\frac{1}{8}\hat{\lambda}_4 - \frac{5}{24}\hat{\lambda}_3^2\right] - \frac{\mu}{\hat{w}_c^2} - \frac{\hat{w}_c}{\hat{t}_c \hat{u}_c} + \frac{c}{\hat{u}_c^2}\left[\frac{\hat{w}_c}{\hat{u}_c} + \frac{1}{2}\hat{w}_c\hat{\lambda}_3\right] \\
\Leftrightarrow B_1(\hat{w}_c) &= -\frac{c}{\hat{u}_c}\left[\frac{1}{8}\hat{\lambda}_4 - \frac{5}{24}\hat{\lambda}_3^2\right] - \frac{\mu}{\hat{w}_c^3} - \frac{1}{\hat{t}_c \hat{u}_c} + \frac{c}{\hat{u}_c^3} + \frac{c\hat{\lambda}_3}{2\hat{u}_c^2}.
\end{aligned}$$

□

B Proof of (12)

Replacing $f_n(y_1)$ in (10) with the first derivative of the appropriate tail probability approximation and applying Theorem 1 yields

$$\begin{aligned}
I_n(c) &= \int_{-\infty}^{c} y_1 f_n(y_1)\mathrm{d}y_1 \\
&= \sqrt{\frac{n}{2\pi}}\int_{-\infty}^{c} \frac{y_1}{\sqrt{\tilde{K}}} e^{-n\tilde{w}^2/2}\left[1 - \frac{1}{n}\left[\frac{\tilde{u}}{\tilde{w}^3} + \tilde{K}^{1/2}\frac{\mathrm{d}}{\mathrm{d}y_1}\frac{1}{\tilde{u}}\right] + \mathcal{O}\left(n^{-1}\right)\right]\mathrm{d}y_1 \\
&= \sqrt{\frac{n}{2\pi}}\int_{-\infty}^{\tilde{w}_c} \frac{\mathrm{d}y_1}{\mathrm{d}\tilde{w}}\frac{y_1}{\sqrt{\tilde{K}}} e^{-n\tilde{w}^2/2}\left[1 - \frac{1}{n}\left[\frac{\tilde{u}}{\tilde{w}^3} + \tilde{K}^{1/2}\frac{\mathrm{d}}{\mathrm{d}y_1}\frac{1}{\tilde{u}}\right] + \mathcal{O}\left(n^{-1}\right)\right]\mathrm{d}\tilde{w} \\
&= \sqrt{\frac{n}{2\pi}}\int_{-\infty}^{\tilde{w}_c} e^{-n\tilde{w}^2/2}\Psi_n(\tilde{w})\mathrm{d}\tilde{w},
\end{aligned}$$

where

$$\begin{aligned}
\Psi_n(\tilde{w}) &\equiv \frac{\mathrm{d}y_1}{\mathrm{d}\tilde{w}}\frac{y_1}{\sqrt{\tilde{K}}}\left[1 - \frac{1}{n}\left[\frac{\tilde{u}}{\tilde{w}^3} + \tilde{K}^{1/2}\frac{\mathrm{d}}{\mathrm{d}y_1}\frac{1}{\tilde{u}}\right] + \mathcal{O}\left(n^{-1}\right)\right] \\
&= \frac{\tilde{w}y_1}{\tilde{u}}\left[1 - \frac{1}{n}\left[\frac{\tilde{u}}{\tilde{w}^3} + \tilde{K}^{1/2}\frac{\mathrm{d}}{\mathrm{d}y_1}\frac{1}{\tilde{u}}\right] + \mathcal{O}\left(n^{-1}\right)\right] \\
&= \psi_0(\tilde{w}) + \frac{\psi_1(\tilde{w})}{n} + \mathcal{O}\left(n^{-1}\right),
\end{aligned} \tag{16}$$

with

$$\psi_0(\tilde{w}) \equiv \frac{\tilde{w}y_1}{\tilde{u}} \quad \text{and} \quad -\psi_1(\tilde{w}) \equiv \frac{y_1}{\tilde{w}^2} + \frac{\tilde{w}y_1\tilde{K}^{1/2}}{\tilde{u}}\frac{\mathrm{d}}{\mathrm{d}y_1}\frac{1}{\tilde{u}}.$$

Again, Equation (16) is in the form of Theorem 1, with the coefficients in the expansion given by

$$B_0(\tilde{w}_c) = \frac{A_0}{\tilde{w}_c} - \frac{c}{\tilde{u}_c}$$

and

$$\begin{aligned}
\tilde{w}_c B_1(\tilde{w}_c) &= A_1 - \psi_1(\tilde{w}_c) + \frac{\mathrm{d}}{\mathrm{d}\tilde{w}_c} B_0(\tilde{w}_c) \\
&= A_1 + \frac{c}{\tilde{w}_c^2} + \frac{\tilde{w}_c c \tilde{K}_c^{1/2}}{\tilde{u}_c} \frac{\mathrm{d}}{\mathrm{d}c} \frac{1}{\tilde{u}_c} + \frac{\mathrm{d}}{\mathrm{d}\tilde{w}_c} \left[\frac{A_0}{\tilde{w}_c} - \frac{c}{\tilde{u}_c} \right] \\
&= A_1 + \frac{c}{\tilde{w}_c^2} + \frac{\tilde{w}_c c \tilde{K}_c^{1/2}}{\tilde{u}_c} \frac{\mathrm{d}}{\mathrm{d}c} \frac{1}{\tilde{u}_c} - \frac{A_0}{\tilde{w}_c^2} - \frac{\mathrm{d}}{\mathrm{d}\tilde{w}_c} \frac{c}{\tilde{u}_c} \\
&= A_1 + \frac{c}{\tilde{w}_c^2} + \frac{\tilde{w}_c c \tilde{K}_c^{1/2}}{\tilde{u}_c} \frac{\mathrm{d}}{\mathrm{d}c} \frac{1}{\tilde{u}_c} - \frac{A_0}{\tilde{w}_c^2} - \frac{\mathrm{d}c}{\mathrm{d}\tilde{w}_c} \frac{\mathrm{d}}{\mathrm{d}c} \frac{c}{\tilde{u}_c} \\
&= A_1 + \frac{c}{\tilde{w}_c^2} + \frac{\tilde{w}_c c \tilde{K}_c^{1/2}}{\tilde{u}_c} \frac{\mathrm{d}}{\mathrm{d}c} \frac{1}{\tilde{u}_c} - \frac{A_0}{\tilde{w}_c^2} - \frac{\tilde{w}_c \tilde{K}_c^{1/2}}{\tilde{u}_c} \left[\frac{1}{\tilde{u}_c} + c \frac{\mathrm{d}}{\mathrm{d}c} \frac{1}{\tilde{u}_c} \right] \\
&= A_1 + \frac{c}{\tilde{w}_c^2} - \frac{A_0}{\tilde{w}_c^2} - \frac{\tilde{w}_c \tilde{K}_c^{1/2}}{\tilde{u}_c^2}.
\end{aligned}$$

□