Article:

Marsh, P.W.N. (1998) Saddlepoint approximations for noncentral quadratic forms. Econometric Theory. pp. 539-559. ISSN 0266-4666

https://doi.org/10.1017/S0266466698145012

Econometric Theory, **14**, 1998, 539–559. Printed in the United States of America.

SADDLEPOINT APPROXIMATIONS FOR NONCENTRAL QUADRATIC FORMS

PATRICK W.N. MARSH
University of York

Many estimators and tests are of the form of a ratio of quadratic forms in normal variables. Excepting a few very special cases little is known about the density or distribution of these ratios, particularly if we allow for noncentrality in the quadratic forms. This paper assumes this generality and derives saddlepoint approximations for this class of statistics. We first derive and prove the existence of an exact inversion based on the joint characteristic function. Then the saddlepoint algorithm is applied and the leading term found, and analytic justification of the asymptotic nature of the approximation is given. As an illustration we consider the calculation of sizes and powers of F-tests, where a new exact result is found.

1. INTRODUCTION

The problem of finding algorithms and explicit formulae for the moments, density, and distributions, or approximations thereof, of a ratio of quadratic forms has been the subject of much attention historically. The reason is simply that a sizable bulk of estimators and test statistics are of this form. Just considering the linear regression model, a battery of specification and misspecification tests conforms to this functional form.

To fix matters we will be interested in a class of statistics, $q(v)$, with

$$q(v) = \frac{v'A_1v}{v'A_2v}, \tag{1.1}$$

with v an N-dimensional random vector, and A_1 and A_2 $N \times N$ constant matrices. In this paper we restrict attention to the class of statistics with v distributed N-dimensional normal, both for computational reasons and simply because it seems most natural, if we have to make some distributional assumption. We will not assume independence of the numerator and denominator, nor that v has zero

I thank Grant Hillier for instructive supervision during my Ph.D., from which this paper originates. Thanks also to John Knight, Jan Podivinsky, Steve Satchell, Irene Valsecchi, participants at the 7th World Congress of the Econometric Society, at Keio University, Tokyo, and two anonymous referees for their comments. Financial support in the form of ESRC awards R00429334128 and R000236627 is gratefully acknowledged. Address correspondence to: Patrick Marsh, Department of Mathematics, University of York, Heslington, York, YO1 5DD, United Kingdom; e-mail: pwnm1@york.ac.uk.

 0266-4666/98 $9.50

mean. The latter may be interpreted as an opportunity to study properties of tests under both the null and the alternative.

In this paper, for such a class of statistics, we will derive saddlepoint approximations for the densities and distributions. Tractable algorithms are presented, and we are able to test both the applicability and accuracy of the approximation by considering size and power calculations for F-tests in the linear regression model. The method is to apply the saddlepoint technique for approximating integrals (e.g., see Bleistein and Handelsman, 1975; De Bruijn, 1961) to obtain an approximation for the density and distribution of $q(v)$. In a statistical context the technique was first exploited by Daniels (1954) and (1956) to obtain approximations for the density of sums of independent variables and for the correlation coefficient. For a full account of the application of the technique, Jensen (1995) contains many results, applications, and a full survey of the literature.

Recently, Lieberman (1994a, 1994b, 1997) has considered saddlepoint-type approximations for statistics similar to $q(v)$. In Lieberman (1994a), approximate moments were found for v having some unspecified distribution, while in Lieberman (1994b), approximations for the density and distribution were found, with v standard normal. Robustness of such approximations was investigated in Lieberman (1997) by allowing v to have some arbitrary, but well-specified, cumulant structure. The approaches there and here are subtly different. Here we begin with the joint characteristic function, rather than the moment generating function (m.g.f.), and obtain, by suitable deformation of the parameter and contour of integration, an inversion dual to that of Geary (1944). We do so only because the implicit assumption of the existence of the joint m.g.f. seems unreasonable, particularly because we are interested in a class of statistics where often it will not (see, e.g., Johnson, Kotz, and Balakrishnan 1995). However, as will become clear later, on application of the saddlepoint technique, the two approaches are essentially identical. This observation becomes apparent because of the approach taken here.

Even restricting attention to normality, only in the most special cases are the exact densities and distributions known, i.e., when A_1 and A_2 are symmetric and idempotent (leading to standard, noncentral, and doubly noncentral F-statistics). For instance consider estimates of autoregressive parameters, with or without exogenous regressors, corresponding to the central or noncentral cases. For these statistics we invariably rely upon approximation techniques, such as here, numerical methods such as Imhof's (1961) procedure, or tedious Monte Carlo studies. For a much fuller account of the properties of statistics of the form given in (1.1), for instance for special cases or alternative approximations in more general cases, see Johnson et al. (1995), for example, Chapters 27 through 32 and references therein. In the context of the linear model, Koerts and Abrahamse (1969) investigated the distribution of ratios of such quadratic forms.

The plan for the remainder of the paper is as follows. Section 2.1 derives an appropriate inversion for the density based on the characteristic function, and 2.2 applies the properties of the statistic $q(v)$ to obtain an exact inversion for the density, for which we are able to prove existence. Section 3.1 gives the application of the saddlepoint algorithm, and in particular the leading term approximation. We then derive some limited analytic properties of the asymptotic nature of the approximation in 3.2. Sections 4.1 and 4.2 derive analogous results for the distribution of the statistic. Section 5 then applies the approximations to the calculation of sizes and powers of F-tests and demonstrates that approximate size calculations are in fact exact, whereas those for the powers are of high accuracy. Concluding remarks are contained in Section 6, and the Appendixes contain some technical derivations.

2. INVERSION FORMULA FOR THE DENSITY

2.1. Preliminaries

As we have noted, because the quadratic forms in (1.1) are noncentral and possibly dependent, and further because we do not assume the existence of the joint m.g.f., Geary's (1944) inversion formula is not directly applicable. Instead we derive a dual inversion, based on the joint characteristic function. This inversion is seen to be of the form of those given by Gurland (1948). We define our statistic as

$$q(v) = \frac{v'A_1 v}{v'A_2 v} = \frac{b_1}{b_2}$$

and the joint characteristic function of b_1 and b_2 as $\chi(\theta_1, \theta_2)$, and we formally make the following assumption.

Assumption 1.

(i) $v \sim N(\mu, I_N)$.
(ii) A_1 and A_2 are symmetric and A_2 is positive definite.
(iii) The function $\chi(\theta_1, \theta_2)$ exists and is convergent in a strip of positive width around the imaginary axes, i.e.,

$$\chi(\tau_1 + i\omega_1, \tau_2 + i\omega_2) < \infty,$$

for some $\tau_1, \tau_2 \in \mathcal{R}^1$.

Arising from Assumption 1 we note the following. First, that v has an identity covariance matrix is not restrictive, because we can always define a $v^* = \Sigma^{-1/2} v$, if $v \sim N(\mu, \Sigma)$. Second, because A_2 is positive definite then

$$\Pr(b_2 \leq 0) = \Pr(v'A_2 v \leq 0) = 0,$$

by continuity of densities. Finally part (iii) assumes that $\chi(\cdot, \cdot)$ remains convergent after deformation in the real direction of the complex plane. Under Assumption 1, and denoting the density of $q(v)$ at a point q as $g(q)$, we have the following lemma.

LEMMA 1. *The exact density of $q(v)$ is obtained by the inversion*

$$g(q) = \frac{1}{2\pi i}\int_{\tau-i\infty}^{\tau+i\infty} \frac{\partial\chi(\theta_1,\theta_2)}{\partial\theta_2}\bigg|_{\theta_2=-q\theta_1} d\theta_1, \tag{2.1}$$

for some $\tau \le \tau_1$ of Assumption 1.

Proof. Lemma 1 follows from a similar result found in Daniels (1956), based on the joint m.g.f. ■

In the following subsection we will apply inversion (2.1) to the form of statistics given in (1.1). The aim is to write the inversion entirely in functions of the fundamental properties of $q(v)$, that is, the matrices A_1 and A_2 and the noncentrality parameter μ.

2.2. The Inversion Formula

Notice that in (2.1) the function $\chi(\cdot,\cdot)$ is no longer the characteristic function itself, but a deformation in the real direction of the complex planes of θ_1 and θ_2. More generally (2.1) is the complex Laplace transform of the joint density of b_1 and b_2.

LEMMA 2. *The Laplace transform $\chi(\theta_1,\theta_2)$, under Assumption* 1, *is given by*

$$\chi(\theta_1,\theta_2) = \exp\{-\lambda\}\exp\left\{\frac{\mu' D^{-1}\mu}{2}\right\}|D|^{-1/2}, \tag{2.2}$$

where $\lambda = \mu'\mu/2$ and $D = [I_N - 2\theta_1 A_1 - 2\theta_2 A_2]$.

Proof. Derivation of Lemma 2 follows from Abadir and Larsson (1996), Theorem 2.1, and existence for θ_1 and θ_2 complex is guaranteed by Ingham (1933), whereas for θ_1 and θ_2 real we also require that D is positive definite. As a note, Assumption 1(iii) is automatically satisfied when θ_1 and θ_2 are complex. ■

Proceeding, we utilize the joint characteristic function to obtain an exact inversion for all statistics given by (1.1) and Assumption 1. What we require is the derivative of (2.2) with respect to θ_2 and then evaluate at the set of points $(\theta_1, -q\theta_1)$. Finally we check for existence of the resulting inversion, on substitution into (2.1). Thus we have the following result.

THEOREM 1.

(i) *The exact density of $q(v)$ at q is*

$$g(q) = \frac{\exp\{-\lambda\}}{2\pi i} \times \int_{\tau-i\infty}^{\tau+i\infty} \exp\left\{\frac{\mu' G^{-1}\mu}{2}\right\}|G|^{-1/2}[\mathrm{Tr}[G^{-1}A_2] + \mu'(G^{-1}A_2G^{-1})\mu]\, d\theta_1, \tag{2.3}$$

with $G = I_N - 2\theta_1 F$ and $F = [A_1 - qA_2]$.

(ii) *Inversion* (2.3) *exists and is continuous everywhere over* θ_1.

Proof. Part (i) is derived by evaluating the derivative of $\chi(\theta_1,\theta_2)$ with respect to θ_2 at $\theta_2 = -q\theta_1$.

For part (ii), supposing θ_1 complex, we make the following points (again the argument follows that of Ingham, 1933).

(i) Because F is real and symmetric its eigenvalues are real, and so the eigenvalues of G are continuous in θ_1. Hence $|G|$ is continuous in θ_1.
(ii) Also, because $\lim_{\theta_1 \to 0}|G|$ is bounded away from zero and by the continuity, above, $|G|$ is never 0.
(iii) As a consequence of (ii) we ensure the existence of both $|G|^{-1/2}$ and G^{-1}. Further because all other terms are constants, for a given statistic, then inversion (2.3) exists for all θ_1 and q as defined in Assumption 1. ■

Because a convergent series representation of the density is as yet unobtainable, in the proceeding section we demonstrate the applicability of a leading term saddlepoint approximation.

3. SADDLEPOINT APPROXIMATIONS

3.1. The Leading Term

In this section we outline how inversion formula (2.3) is of a form for which a saddlepoint approximation is easily obtainable. In particular, through an appropriate expansion of the kernel and subsequent transformation, the method of Laplace (see Barndorff-Nielsen and Cox, 1989, Ch. 3; De Bruijn, 1961, Ch. 4) delivers a leading term saddlepoint approximation, relatively tractable in form. We can rewrite (2.3) so that

$$g(q) = \frac{\exp\{-\lambda\}}{2\pi i}\int_{\tau-i\infty}^{\tau+i\infty} \exp\{P(\theta_1)\}Q(\theta_1)\,d\theta_1, \tag{3.1}$$

where we have defined

$$P(\theta_1) = P(\tau,\omega_1) = \tfrac{1}{2}(\mathrm{Tr}[G^{-1}S] - \ln|G|)$$

and

$$Q(\theta_1) = P(\tau,\omega_1) = \mathrm{Tr}[G^{-1}A_2(G^{-1}S + I)],$$

where above $S = \mu\mu'$ is a positive semidefinite matrix, possibly zero only through $\mu = 0$, with G and A_2 defined before, and noting $\theta_1 = \tau + i\omega_1$.

Inversion (3.1) is of the form considered by De Bruijn (1961, Ch. 4), and hence the leading term saddlepoint approximation is simply

$$\hat{g}(q) = \frac{\exp\{-\lambda\}\exp\{P(\hat{\tau})\}Q(\hat{\tau})}{\sqrt{2\pi|P''(\hat{\tau})|}}, \tag{3.2}$$

where $P(\hat{\tau}) = P(\hat{\tau},0)$ and $Q(\hat{\tau}) = Q(\hat{\tau},0)$ and the saddlepoint $\hat{\tau}$ is defined by

$$P'(\hat{\tau}) = P(\tau,\omega_1)|_{\tau=\hat{\tau},\omega_1=0} = 0. \tag{3.3}$$

In summary, we have basically transformed the integration problem, equation (3.1), into the optimization problem (3.3). Strictly speaking, the saddlepoint defined by (3.3) is a saddlepoint only of P and consequently $\exp\{P\}$, not the kernel itself as in Daniels (1954, 1987). However, the technique suffices in this case because the exponent dominates the remaining term in (3.1).

We can apply approximation (3.2) to our inversion formula (2.3), noting that

$$P'(\tau) = \mathrm{Tr}[G^{-1}F(G^{-1}S + I)]$$

and

$$P''(\tau) = 2\,\mathrm{Tr}[(G^{-1}F)^2(2G^{-1}S + I)]. \tag{3.4}$$

Denoting $\hat{G}$ as G evaluated at the point $\theta_1 = \hat{\tau}$, then the leading term saddlepoint approximation to the density of a ratio of dependent, noncentral quadratic forms is

$$\hat{g}(q) = \frac{\exp\{-\lambda\}etr\{\hat{G}^{-1}S\}\mathrm{Tr}[\hat{G}^{-1}A_2(\hat{G}^{-1}A_2(\hat{G}^{-1}S + I)]}{|\hat{G}|\sqrt{4\pi|\mathrm{Tr}[(\hat{G}^{-1}F)^2(2\hat{G}^{-1}S + I)]|}}, \tag{3.5}$$

and the saddlepoint $\hat{\tau}$ is defined by

$$\mathrm{Tr}(G^{-1}F(G^{-1}S + I)) = 0. \tag{3.6}$$

Unfortunately, for computational purposes (3.6) is of little use, because symbolic matrices are computationally expensive to invert, however, in Appendix A, we demonstrate that the function $P(\theta_1)$ may be written as

$$P(\theta_1) = \frac{1}{2}\left[\sum_{i=1}^{N}(u_i)^2(1 - 2\theta_1 f_i)^{-1} - \sum_{i=1}^{N}\ln(1 - 2\theta_1 f_i)\right], \tag{3.7}$$

where the f_i are the ordered eigenvalues of F, the u_i are the elements of the vector $u = R\mu$, and R diagonalizes G. Differentiating (3.7) directly then gives

$$P'(\theta_1) = \left[\sum_{i=1}^{N}(u_i)^2(1 - 2\theta_1 f_i)^{-2}f_i + \sum_{i=1}^{N}(1 - 2\theta_1 f_i)^{-1}f_i\right], \tag{3.8}$$

whose roots can be found by most numerical packages. In particular, there are $N-1$ solutions to (3.8), each lying between the asymptotes $(2f_i)^{-1}$. Because $\tilde{g}(q)$ is real only for

$$\hat{\tau} \in ((2\max_i(f_i))^{-1}, (2\min_i(f_i))^{-1}), \tag{3.9}$$

(see Daniels, 1954), the saddlepoint is found by employing a line search in (3.8), constrained to this region. It has been found in applications of (3.5) that the roots of (3.8) tend to be more numerically stable than those obtained from (3.6), hence the suggested use of this form. Calculations in this paper were performed with Mathematica™ (see Wolfram, 1991); for other packages, the roots of (3.6) may turn out to be more stable.

As a note, because the saddlepoint is constrained to the region defined in (3.9), then obviously $\hat{G}^{-1}$ exists. Hence for the set of points q and $\theta_1(q) = \hat{\tau}$, inversion (2.3) exists, independently of the argument of Theorem 1. As a consequence, as far as the saddlepoint technique is concerned, the joint m.g.f. approach will also yield a valid approximation.

Given approximation (3.5), one pertinent question is, to what extent is it a higher-order asymptotic expansion? We address this issue next.

3.2. Asymptotic Nature in the Tails

Here we consider the issue of examining the asymptotic nature of the approximation. We note immediately that the form of approximation (3.5) differs from those given in Daniels (1954, 1956). There, the function we define as $P(\theta_1)$ is obviously $O(N)$, and hence the asymptotic nature of those series was simple, i.e., in powers of N^{-1}. In our case $P(\theta_1)$ and $Q(\theta_1)$ depend more subtly with N, and to obtain analytic results we examine only the far tails of the distribution. We note that for extreme values of q the numerical value of the exact density is zero; our aim here is to determine the rate of convergence in N, of the approximation to this value.

First we demonstrate the limiting relation between q and $\hat{\tau}$ (i.e., between the evaluation point and the saddlepoint), which we present as the following lemma.

LEMMA 3. *For $\mu \neq 0$ and $|q|$ sufficiently large, there exists the unique saddlepoint, not necessarily finite, given by*

$$\hat{\tau}(q) = \frac{\mathrm{Tr}[(S+I)F]}{2\,\mathrm{Tr}[F^2]}, \tag{3.10}$$

where the dependence of $\hat{\tau}$ (the saddlepoint) upon q (the point at which the approximation is evaluated) is made explicit in the limit.

Proof. The proof of Lemma 3 is given in Appendix B.

Utilizing the previous lemma, we are able to examine asymptotic behavior of the approximations in the very far tails of the distribution. Upon full expansion of (3.1) and ignoring constants, a typical term is

$$h_{j,m}(\hat{\tau}) = \left[\frac{Q^{(j)}(\hat{\tau})}{Q(\hat{\tau})\{P''(\hat{\tau})\}^{j/2}}\right]\left[\sum_{k=3}^{\infty}\frac{P^{(k)}(\hat{\tau})}{\{P''(\hat{\tau})\}^{k/2}}\right]^m, \tag{3.11}$$

$$j,m = 0,1,2,\ldots,$$

where $Q^{(j)}(\hat{\tau})(P^{(k)}(\hat{\tau}))$ is the jth (kth) derivative of $Q(P)$ at $\theta_1 = \hat{\tau}$. Further, we can show that

$$Q^{(j)}(\hat{\tau}) \propto \mathrm{Tr}[(\hat{G}^{-1}F)^j \hat{G}^{-1} A_2 (j\hat{G}^{-1}S + I)],$$
$$P^{(k)}(\hat{\tau}) \propto \mathrm{Tr}[(\hat{G}^{-1}F)^k (k\hat{G}^{-1}S + I)], \tag{3.12}$$

verified by differentiation and thence by induction from the $j, k = 1$ case. In Appendix C we demonstrate that $|q| \to \infty$,

$$\lim_{|q|\to\infty} \left. \begin{aligned} Q^{(j)}(\hat{\tau}(q)) &= O(N) \\ P^{(k)}(\hat{\tau}(q)) &= O(N) \end{aligned} \right\} \quad \forall j, k, \tag{3.13}$$

and consequently we have

$$h_{j,m}(\hat{\tau}(q)) \propto \frac{1}{N^{j/2}} \left[\frac{1}{N^{k/2-1}} \right]^m. \tag{3.14}$$

Because odd powers disappear upon integration the leading term has order of error $O(N^{-1})$, the leading term plus first correction $O(N^{-2})$, and so on.

Having derived the leading term for the density we now turn our attention to the distribution function of $q(v)$.

4. APPROXIMATING THE DISTRIBUTION

4.1. An Inversion

Here we consider the problem of finding an approximation for the distribution function of $q(v)$, i.e., $G(q) = \Pr\{q(v) \le q\}$, or if $q(v)$ is in particular some test statistic, the tail area probability, $\Pr\{q(v) > q\} = 1 - G(q)$. We can proceed in two possible ways. First, we could simply numerically integrate approximation (3.2) over the relevant range. Alternatively, we can seek to provide an analytic solution to this problem, i.e., find an approximation for the distribution analogous to that found for the density. Clearly there are benefits to finding an analytic rather than a numeric approximation, particularly if the accuracy of the approximation remains good, for instance, if $q(v)$ is a test statistic then finding critical values will prove far simpler in this case.

To begin we require an inversion for the distribution; the following lemma, simply derived from Lemma 1, gives the result.

LEMMA 4. *Under the conditions of Assumption* 1, *and denoting the distribution of $q(v)$ as $G(q)$, then*

$$G(q) = 1 - \frac{\exp\{-\lambda\}}{2\pi i} \int_{\tau - i\infty}^{\tau + i\infty} \frac{\exp\{P(\theta_1)\}}{\theta_1} \, d\theta_1. \tag{4.1}$$

Because inversion (4.1) is essentially similar in nature to the inversion formula for the density, equation (3.1), but with $Q(\theta_1)$ replaced by θ_1^{-1}, then an attempt to apply the saddlepoint technique to approximate it seems natural.

4.2. Saddlepoint Approximation for the Distribution

Simple application of the saddlepoint method, for inversion (4.1), breaks down because the range of integration will pass through both a pole (at $\theta_1 = 0$) and a saddlepoint (at $\theta_1 = \hat{\tau}$). To obtain an approximation reliable over the entire range we show here that the integral in (4.1) may be transformed to a form that yields the Lugannani and Rice (1980) approximation. Again we consider only the integral in (4.1), defined as

$$I_q = \frac{1}{2\pi i} \int_{\tau - i\infty}^{\tau + i\infty} \frac{\exp\{P(\theta_1)\}}{\theta_1} \, d\theta_1 \tag{4.2}$$

such that

$$G(q) = 1 - \exp\{-\lambda\} I_q.$$

Following Daniels (1987), the saddlepoint approximation to (4.2), on isolation of the pole, is the Lugannani and Rice formula:

$$\hat{I}_q = 1 - \Phi(\hat{p}) + \phi(\hat{p}) \left[\frac{1}{\hat{r}} - \frac{1}{\hat{p}} \right]. \tag{4.3}$$

Here, $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution functions, respectively, $\hat{p}$ and $\hat{r}$ are real and defined by

$$\hat{p} = \operatorname{sign}(\hat{\tau}) \sqrt{2|P(\hat{\tau})|} \qquad \text{and} \qquad \hat{r} = \hat{\tau} \sqrt{2|P''(\hat{\tau})|} \tag{4.4}$$

and once again $\hat{\tau}$ defined by the relevant root of (3.8). Alternatively, Jensen (1995, Ch. 3), suggested the alternative form

$$\hat{I}_q = \Phi\left(-\hat{p} + \frac{1}{\hat{p}} \log\left[\frac{\hat{p}}{\hat{r}} \right] \right)$$

for the saddlepoint approximation for (4.2).

If we return to our statistic $q(v)$, noting expressions found for $P(\hat{\tau})$ and $P''(\hat{\tau})$ previously and substituting into $\hat{p}$ and $\hat{r}$ gives us that the saddlepoint approximation for the distribution function of $q(v)$ is given by

$$\hat{G}(q) = (1 - e^{-\lambda}) + e^{-\lambda} \left(\Phi(\hat{p}) - \phi(\hat{p}) \left[\frac{1}{\hat{r}} - \frac{1}{\hat{p}} \right] \right) \tag{4.5a}$$

(Daniels, 1987) and

$$\hat{G}(q) = (1 - e^{-\lambda}) + e^{-\lambda}\Phi\left(\hat{p} - \frac{1}{\hat{p}}\log\left[\frac{\hat{p}}{\hat{r}}\right]\right) \tag{4.5b}$$

(Jensen, 1995), with $\hat{p}$ and $\hat{r}$ given by

$$\hat{p} = \mathrm{sign}(\hat{\tau})\sqrt{|\mathrm{Tr}[\hat{G}^{-1}S] - \ln|\hat{G}||},$$
$$\hat{r} = 2\hat{\tau}\sqrt{|\mathrm{Tr}[(\hat{G}^{-1}F)^2(2\hat{G}^{-1}S + I)]|},$$

and λ is defined previously. As a note, the approximations in (4.5) may be interpreted as an exponential mixture of Dirac (unit point mass on the first term) and normal distributions, where the mixture depends only upon the noncentrality parameter λ, not upon the matrices A_1 and A_2, whose role is limited to the argument of the normal.[1]

In (3.7) and (4.5) we have leading term saddlepoint approximations for the density and distribution of $q(v)$, under Assumption 1. In the following section we will apply these results to the calculation of sizes and powers of F-tests in a linear regression.

5. THE SIZE AND POWER OF *F*-TESTS

Take the linear regression model, defined by

$$Y = X\beta + \varepsilon,$$

where $\varepsilon \sim N(0, \sigma^2 I)$, with $Y = \{y_1, y_2, \ldots, y_N\}$, X is an $N \times k$ matrix of regressors of full rank and β is a $k \times 1$ vector of coefficients, and consider the problem of testing the simple null hypothesis:

$$H_0: \beta = 0 \qquad \text{against} \qquad H_1: \beta = \beta^*.$$

Because in general there exists no uniformly most powerful test for this problem, the F-test, which maximizes average power, is most often employed. The test statistic, in our notation, has the form

$$F = \frac{N-k}{k} q(Y),$$

$$q(Y) = \frac{Y'P_X Y}{Y'M_X Y} = \frac{q_1}{q_2},$$

where $P_X = X(X'X)^{-1}X'$ and $M_X = I - P_X$. We may also note that $q(Y)$ may be interpreted as the likelihood ratio test, for this hypothesis.

We utilize the results of the previous sections to calculate approximations for the density and distribution of $q(v)$ under both the null (the central case) and alternative (the noncentral case). We then compare the approximation for the power of the F-test with the exact result to indicate the numerical accuracy of the approximation.

5.1. The Null Hypothesis

Under the null hypothesis $q(v)$ is a ratio of independent central quadratic forms (i.e., $\mu = X\beta = 0$). The Laplace transform, yielding the joint characteristic function and m.g.f., is given by equation (2.1) with $\mu = 0$:

$$\chi(\theta_1,\theta_2) = |I - 2\theta_1 P_X - 2\theta_2 M_X|^{-1/2},$$

and noting $M_X = I - P_X$, where P_X is idempotent, then we obtain the well-known result

$$\chi(\theta_1,\theta_2) = (1-2\theta_1)^{-k/2}(1-2\theta_2)^{-(N-k)/2}. \tag{5.1}$$

To calculate the appropriate inversion formula, we may either directly differentiate (5.1) or the more general (2.2); the latter yields

$$\left.\frac{\partial\chi(\theta_1,\theta_2)}{\partial\theta_2}\right|_{\theta_2=-q\theta_1} = |aI + bP_X|^{-1/2}\,\mathrm{Tr}[(aI+bP_X)^{-1}M_X], \tag{5.2}$$

where $a = (1+2q\theta_1)$ and $b = -2\theta_1(1+q)$. After some manipulation we find the inversion formula for $q(v)$ under the null is

$$g(q) = \frac{(N-k)}{2\pi i}\int_{\tau-i\infty}^{\tau+i\infty}(1-2\theta_1)^{-k/2}(1+2q\theta_1)^{-(N-k)/2-1}\,d\theta_1. \tag{5.3}$$

In the notation of this paper we write the inversion formula in terms of the functions P and Q, as before, where for this example

$$P(\theta_1) = -\tfrac{1}{2}\ln[|aI+bP_X|] \quad \text{and} \quad Q(\theta_1) = \mathrm{Tr}[(aI+bP_X)^{-1}M_X],$$

and so application of (3.2) is relatively straightforward, and in particular the relevant derivatives of P are

$$P'(\theta_1) = \mathrm{Tr}[(aI+bP_X)^{-1}(P_X - qM_X)] = \frac{k}{(1-2\theta_1)} - \frac{q(N-k)}{(1+2q\theta_1)},$$

$$P''(\theta_1) = 2\,\mathrm{Tr}[((aI+bP_X)^{-1}(P_X-qM_X))^2] = 2\left(\frac{k}{(1-2\theta_1)^2} - \frac{q^2(N-k)}{(1+2q\theta_1)^2}\right). \tag{5.4}$$

The saddlepoint for this example is defined by $P'(\theta_1) = 0$, which yields the unique saddlepoint

$$\hat{\tau} = \frac{q(N-k)-k}{2qN}. \tag{5.5}$$

Substitution of the saddlepoint into P and Q of equation (5.4) and some manipulation give us the leading term saddlepoint approximation for $q(v)$ under the null as

$$\hat{g}(q) = \frac{k^{-(k+1)/2}(N-k)^{-(N-k+1)/2}(1+q)^{-N/2}q^{-(k+2)/2}}{2\sqrt{\pi}N^{(N+1)/2}}. \tag{5.6}$$

Moreover, as a direct consequence of approximation (5.6) we obtain the following result.

Result. The leading term saddlepoint approximation for $q(v)$ under the null is exact, up to a normalizing constant. If we write the approximation as

$$\hat{g}(q) = a(k,N)(1+q)^{-N/2}q^{-(k+2)/2},$$

with $a(k,N)$ a constant, then renormalizing in the sense of Daniels (1954) gives the exact density, because

$$\int_q (1+q)^{-N/2}q^{-(k+2)/2}\,dq = B\left(\frac{k}{2},\frac{N-k}{2}\right),$$

the beta function. This adds to the list of cases for which the saddlepoint technique yields exact distributional results (see Daniels, 1980), although the result here is different in nature, because the statistic is not a sum of independent and identically distributed (i.i.d.) variables. ■

Now because the saddlepoint technique yields an exact density for $q(v)$ under the null, then the simple transformation to the central F-statistic will yield exact densities. Consequently, we do not require the approximation for the distribution, because integration of (5.6) is straightforward.

5.2. The Alternative Hypothesis

Under the alternative, $\beta = \beta^*$, $q(v)$ is a ratio of a noncentral and a central quadratic form, again independent. Again we utilize equation (2.1). In our example this becomes

$$\chi(\theta_1,\theta_2) = e^{-\lambda}\exp\left\{\frac{\beta^{*\prime}X'(I-2\theta_1P_X-2\theta_2M_X)^{-1}X\beta^*}{2}\right\} \times |I-2\theta_1P_X-2\theta_2M_X|^{-1/2}. \tag{5.7}$$

Proceeding as in the central case we find

$$|I-2\theta_1P_X-2\theta_2M_X|^{-1/2} = (1-2\theta_1)^{-k/2}(1-2\theta_2)^{-(N-k)/2},$$

$$(I-2\theta_1P_X-2\theta_2M_X)^{-1} = \frac{1}{(1-2\theta_2)}\left(I+\frac{2(\theta_1-\theta_2)}{1-2\theta_2}P_X\right),$$

which gives us the well-known result

$$\chi(\theta_1,\theta_2) = \exp\left\{\frac{2\lambda\theta_1}{1-2\theta_1}\right\}(1-2\theta_1)^{-k/2}(1-2\theta_2)^{-(N-k)/2}. \tag{5.8}$$

The relevant inversion formula under the alternative can be obtained again by directly differentiating (5.8), or equation (2.2), yielding

$$\left.\frac{\partial\chi(\theta_1,\theta_2)}{\partial\theta_2}\right|_{\theta_2=-q\theta_1} = (N-k)\exp\left\{\frac{\lambda\theta_1}{(1-2\theta_1)}\right\}(1-2\theta_1)^{-(N-k)/2}(1+2q\theta_2)^{-k/2-1}, \tag{5.9}$$

which we simply substitute into the inversion. After further manipulation we have the saddlepoint approximation for the density of $q(v)$, in the noncentral case, as

$$\hat{g}(q) = \frac{(N-k)\exp\left\{\frac{\lambda\hat{\tau}}{1-2\hat{\tau}}\right\}(1+2q\hat{\tau})^{-(N-k)/2-1}(1-2\hat{\tau})^{-k/2}}{\sqrt{4\pi\left|\frac{2\lambda}{(1-2\hat{\tau})^3}+\frac{k}{(1-2\hat{\tau})^2}\right|}}, \tag{5.10}$$

with the saddlepoint defined by substitution into (3.8). For the alternative the approximation is no longer exact, and an approximation for the distribution (and hence powers) follows from substitution into equation (4.5a). We shall not report the detail here.

Because the application of the approximations has proved relatively straightforward, if somewhat messy in the noncentral case, for the remainder of this example we will concentrate upon a comparison of their accuracy with "exact" results. First we report tables comparing approximate versus exact tail area probabilities for the noncentral F distribution with $k = 1$, $N = 10$ or 20, and $\lambda = 1$ or 2. Tables 1 and 2 give tail area probabilities for two noncentral F-statistics, i.e., the F-test under the alternative, corresponding to the linear regression with a single regressor, and sample sizes 10 and 20. Table 1 assumes the alternative is such that $\beta^{*\prime}X'X\beta^* = 1$, whereas Table 2 assumes an alternative such that $\beta^{*\prime}X'X\beta^* = 2$. The accuracy of the approximation for these tail area probabilities is exceedingly good, always within 5% or so, even at a sample size of 10, although we can see a slight, but noticeable, improvement as the sample size increases to 20, giving some further justification to the asymptotic nature of the approximation. In fact this approximation clearly outperforms the limiting approximation, i.e.,

$$F'_{(k,N-k,2\lambda)} \to \chi'^2(k,2\lambda) \quad \text{as } N \to \infty,$$

even though this limiting distribution is itself little simpler than the exact, although details of this are not given.

TABLE 1. Noncentrality parameter, $\lambda = 0.5$

	$\Pr\{F_{(1,9,1)} > q^*\}$		$\Pr\{F_{(1,19,1)} > q^*\}$	
q^*	Exact	Approximation	Exact	Approximation
5	0.129275	0.134001	0.150660	0.144903
5.5	0.110217	0.113113	0.131793	0.124794
6	0.094202	0.095699	0.115736	0.107909
6.5	0.080707	0.081152	0.102004	0.093669
7	0.069303	0.068974	0.090209	0.081610
7.5	0.059643	0.058757	0.080035	0.071355
8	0.051440	0.050166	0.071226	0.062599
8.5	0.044457	0.042926	0.063570	0.055095
9	0.038500	0.036810	0.056894	0.048640
9.5	0.033408	0.031633	0.051053	0.043067

Tables 3 and 4 tabulate power functions for testing the simple hypothesis mentioned previously, at the 5% significance level. The values given in the table are the tail area probabilities from the critical values given by fixing the size of the test at 5% and varying the alternative β^*, such that λ^* takes the values given. These critical values are given by

Table 3: $k = 1$; $N = 10$: $F_{(1,9)}^{.05} = 5.12$; $N = 20$: $F_{(1,19)}^{.05} = 4.38$,

Table 4: $k = 3$; $N = 10$: $F_{(3,9)}^{.05} = 3.86$; $N = 20$: $F_{(3,19)}^{.05} = 3.13$.

TABLE 2. Noncentrality parameter, $\lambda = 1$

	$\Pr\{F_{(1,9,2)} > q^*\}$		$\Pr\{F_{(1,19,2)} > q^*\}$	
q^*	Exact	Approximation	Exact	Approximation
5	0.228381	0.234579	0.251153	0.242923
5.5	0.200162	0.204274	0.224330	0.214261
6	0.175571	0.177930	0.200815	0.189371
6.5	0.154137	0.155062	0.180156	0.167733
7	0.135448	0.135225	0.161966	0.148891
7.5	0.119142	0.118022	0.145917	0.132456
8	0.104907	0.103103	0.131725	0.118092
8.5	0.092472	0.090161	0.119149	0.105512
9	0.081599	0.078928	0.107981	0.094472
9.5	0.072085	0.069172	0.098044	0.084762

Table 3. Number of regressors, $k = 1$

	Power ($N = 10$)		Power ($N = 20$)	
$2\lambda^*$	Exact	Approximation	Exact	Approximation
0.1	0.068609	0.069257	0.070964	0.081296
0.3	0.106821	0.103103	0.114035	0.122003
0.5	0.145846	0.139860	0.158127	0.166110
0.7	0.185269	0.177861	0.202660	0.211083
0.9	0.224735	0.216413	0.247140	0.256088
1.1	0.263939	0.255083	0.291155	0.300547
1.3	0.302622	0.293525	0.334361	0.344027
1.5	0.340569	0.331458	0.376478	0.386202
1.7	0.377603	0.368655	0.417285	0.426837
1.9	0.413578	0.404936	0.456609	0.465763

Again we notice the high accuracy of the approximation; the approximation always lies within 5% of the exact over the range considered here, and there is slight improvement asymptotically.

6. CONCLUDING REMARKS

The paper has established the existence of inversion formulae for the density and distribution of a ratio of noncentral quadratic forms in normal variables. The saddlepoint approximation to these inversions has been found, and we have il-

Table 4. Number of regressors, $k = 3$

	Power ($N = 10$)		Power ($N = 20$)	
$2\lambda^*$	Exact	Approximation	Exact	Approximation
0.1	0.058138	0.053187	0.059683	0.068165
0.3	0.075032	0.065629	0.080503	0.085004
0.5	0.092868	0.080508	0.102770	0.105736
0.7	0.111534	0.096708	0.126289	0.128370
0.9	0.130924	0.113917	0.150867	0.152354
1.1	0.150936	0.131967	0.176321	0.177379
1.3	0.171471	0.150734	0.202472	0.203210
1.5	0.192434	0.170111	0.229155	0.229646
1.7	0.213738	0.190003	0.256211	0.256508
1.9	0.235298	0.210320	0.283493	0.283632

lustrated the procedure with a simple example. Using a joint characteristic function–based approach, not only do we establish the existence of exact inversions to which the technique may be applied, but we also demonstrate that by defining the saddlepoint suitably, the approximations can equally be obtained via an inversion of the joint m.g.f.

We have also considered the application of the approximations to the simplest cases, the central and noncentral F-statistics. In particular we obtain a further example of where the technique delivers an exact result, the central case, and demonstrate the accuracy of the approximation in the noncentral case. Other possible applications for the results obtained here are numerous. As examples we could consider tests for an autoregressive parameter in the linear regression, based on its ordinary least squares (OLS) estimate, or for likelihood (variance) ratio tests for non-nested models. In both cases the exact densities are unknown, but the statistics conform to our statistic and assumptions. However, the issues involved in such applications quickly become computational in nature, i.e., the issue of finding a solution to the saddlepoint defining equation. For these cases an approximate, rather than exact, saddlepoint may prove more useful.

Whereas explicit analytic results for the asymptotic nature of the approximation are difficult to obtain, we have given some justification to the claim of a higher-order asymptotic expansion, although only in the far tails. However, perhaps a more relevant criterion is that the approximations seem competitive with exact results and also significantly improve upon the first-order result.

NOTE

1. Thanks are due to an anonymous referee for this interpretation.

REFERENCES

Abadir, K.M. & R. Larsson (1996) The joint moment generating function of quadratic forms in multivariate autoregressive series. *Econometric Theory* 12, 682–704.

Barndorff-Nielsen, O.E. & D.R. Cox (1989) *Asymptotic Techniques for Use in Statistics.* London: Chapman and Hall Ltd.

Bleistein, N. & N. Handelsman (1975) *Asymptotic Expansions of Integrals.* New York: Holt, Reinhart and Winston.

De Bruijn, N.G. (1961) *Asymptotic Methods in Analysis.* Amsterdam: North-Holland.

Daniels, H.E. (1954) Saddlepoint approximations in statistics. *Annals of Mathematics and Statistics* 25, 631–650.

Daniels, H.E. (1956) The approximate distribution of the serial correlation coefficient. *Biometrika* 43, 169–185.

Daniels, H.E. (1980) Exact saddlepoint approximations. *Biometrika* 67, 53–58.

Daniels, H.E. (1987) Tail probability approximations. *International Statistical Review* 55, 37–48.

Geary, R.C. (1944) Extension of a theorem by Harold Cramer on the frequency distribution of two variables. *Journal of the Royal Statistical Society*, *Series B* 107, 65–67.

Gurland, J. (1948) Inversion formula for the distribution of ratios. *Annals of Mathematical Statistics* 19, 228–237.

Imhof, J.P. (1961) Computing the distribution of quadratic forms in normal variables. *Biometrika* 48, 419–426.

Ingham, A.E. (1933) An integral which appears in statistics. *Proceedings of the Cambridge Philosophical Society* 29, 271–276.
Jensen, J.L. (1995) *Saddlepoint Approximations.* Oxford Statistical Science Series 16. Oxford: Clarendon Press.
Johnson, N., N.L. Kotz, & N. Balakrishnan (1995) *Continuous Univariate Distributions*, vol. 2, 2nd ed. New York: Wiley.
Koerts, J. & A.P.J. Abrahamse (1969) *The Theory and Application of the General Linear Model.* Rotterdam: University Press.
Lieberman, O. (1994a) A Laplace approximation to the moments of a ratio of quadratic forms. *Biometrika* 81, 681–690.
Lieberman, O. (1994b) Saddlepoint approximations to the distribution of a ratio of quadratic forms in normal variables. *Journal of the American Statistical Association* 89, 924–928.
Lieberman, O. (1997) On the effect of non-normality. *Econometric Theory* 13, 52–78.
Lugannani, R. & S. Rice (1980) Saddlepoint approximations for the distribution of the sum of independent random variables. *Advances in Applied Probability* 12, 475–490.
Magnus, J.R. & H. Neudecker (1988) *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Wiley Series in Probability and Statistics. New York: Wiley.
Wolfram, S. (1991) *Mathematica: A System for Doing Mathematics by Computer*, 2nd ed. New York: Addison-Wesley.

APPENDIX A: ON THE SADDLEPOINT DEFINING EQUATION

Returning to equation (3.8), we first note that the trace in $P(\theta_1)$ is equal to the quadratic form

$$w(\theta_1) = \mu'(I - 2\theta_1 F)^{-1}\mu. \tag{A.1}$$

Now let R be the $N \times N$ diagonalizing matrix such that

$$RG^{-1}R' = D(g^-),$$

where $D(g^-)$ is the $N \times N$ diagonal matrix with structure

$$D(g^-) = \begin{pmatrix} g_1^- & 0 & . & 0 \\ 0 & g_2^- & & : \\ : & & . & : \\ 0 & . & . & g_N^- \end{pmatrix}$$

and the g_i^- are the N ordered eigenvalues of G^{-1} such that $g_1^- > g_2^- > \cdots > g_N^-$.

Consequently, the quadratic form $w(\theta_1)$ may be expressed in canonical form as

$$w(\theta_1) = \sum_{i=1}^{N} (u_i)^2 g_i^-. \tag{A.2}$$

Further, because we have that G and hence G^{-1} are of full rank then $g_i^- \neq 0 \; \forall i$. So defining a series $\{g_i\}$ the eigenvalues of G with $g_i^- = (g_i)^{-1}$ and the g_i^- preserve their order. Thus we have the simpler form

$$w(\theta_1) = \sum_{i=1}^{N} (u_i)^2 (g_i)^{-1}.$$

From the definition $G = (I - 2\theta_1 F)$, the eigenvalues of G are

$$g_i = (1 - 2\theta_1 f_i),$$

where the f_i are the N eigenvalues of F, and so we can write $w(\theta_1)$ in terms only of the parameter θ_1 and the eigenvalues of the known matrix, $F = A_1 - qA_2$, i.e.,

$$w(\theta_1) = \sum_{i=1}^{N} (u_i)^2 (1 - 2\theta_1 f_i)^{-1}. \tag{A.3}$$

Now if we return to the remaining term in the definition of $P(\theta_1)$, i.e., $\ln|G|$, the well-known result gives us

$$\ln|G| = \ln|I - 2\theta_1 F| = \sum_{i=1}^{N} \ln(1 - 2\theta_1 f_i), \tag{A.4}$$

where the f_i are the eigenvalues defined and ordered as previously. So in light of (A.4) and (A.2) we can write

$$P(\theta_1) = \frac{1}{2}\left[\sum_{i=1}^{N} (u_i)^2 (1 - 2\theta_1 f_i)^{-1} - \sum_{i=1}^{N} \ln(1 - 2\theta_1 f_i)\right] \tag{A.5}$$

and hence is a function entirely of the eigenvalues of the known constant matrix F and the argument of the inversion formula. ■

APPENDIX B: PROOF OF LEMMA 3

Recall the saddlepoint defining equation:

$$\mathrm{Tr}[G^{-1}F(G^{-1}S + I)] = 0, \tag{B.1}$$

and F and G defined previously. If $|q|$ is sufficiently large then either

F positive definite; $\quad q < 0$,

F negative definite; $\quad q > 0$.

We take $q > 0$ and note that the opposite case follows similarly. Because F is negative definite, write

$$-F = BB',$$

with $B; N \times m$, of rank $m \leq N$. Hence $G = I + 2\theta_1 BB'$, and so

$$\mathrm{Tr}[G^{-1}BB'(G^{-1}S + I)] = 0,$$

or alternatively

$$\mathrm{Tr}[B'G^{-1/2}(G^{-1/2}SG^{-1/2} + I)G^{-1/2}B] = 0. \qquad \textbf{(B.2)}$$

Because $\mu \neq 0$, $S > 0$ and $(G^{-1/2}SG^{-1/2} + I) > 0$, so (B.1) holds for all $\hat{\tau}$ and hence $\hat{G}$ that satisfy

$$(\hat{G}^{-1/2}S\hat{G}^{-1/2} + I)\hat{G}^{-1/2}B = 0. \qquad \textbf{(B.3)}$$

(B.3) follows because, for $W > 0$,

$$\mathrm{Tr}\{X'WX] = 0 \Leftrightarrow WX = 0,$$

as in Magnus and Neudecker (1988, Sect. 12, p. 367). We have shown $\hat{G}^{-1/2}$ exists, and so (B.3) implies

$$(S\hat{G}^{-1} + I)B = 0,$$

which gives us the useful relations

$$\begin{aligned} B &= -S\hat{G}^{-1}B, \\ BB'B &= -S\hat{G}^{-1}BB'B, \\ 2\hat{\tau}BB'B &= -S\hat{G}^{-1}(\hat{G} - I)B = -(B + SB), \\ SB &= -B(I_m + 2\hat{\tau}B'B). \end{aligned} \qquad \textbf{(B.4)}$$

Returning our attention to (B.1) then we also have

$$\begin{aligned} \hat{G}^{-1} &= I - 2\hat{\tau}B[I_m + 2\hat{\tau}B'B]^{-1}B', \\ \hat{G}^{-1}S &= S - 2\hat{\tau}B[I_m + 2\hat{\tau}B'B]^{-1}B'S \end{aligned} \qquad \textbf{(B.5)}$$

and because $(B'S) = (SB)'$ and from (B.4)

$$\hat{G}^{-1}S = S + 2\hat{\tau}BB' \ (= S - 2\hat{\tau}F).$$

So (B.1) may be rewritten as

$$\mathrm{Tr}[(S + 2\hat{\tau}BB' + I)\hat{G}^{-1}F] = 0, \qquad \textbf{(B.6)}$$

and we also have

$$\begin{aligned} S\hat{G}^{-1} &= S - 2\hat{\tau}F \quad \text{(by symmetry)}, \\ 2\hat{\tau}BB'\hat{G}^{-1} &= 2\hat{\tau}B(I_m + 2\hat{\tau}B'B)^{-1}B' \quad \text{(from (B.4) and (B.5))}, \end{aligned} \qquad \textbf{(B.7)}$$

then simply collecting terms in (B.6) and writing $\hat{\tau} = \hat{\tau}(q)$,

$\mathrm{Tr}[(\hat{G}^{-1}S + I)\hat{G}^{-1}F] = 0$

gives

$\mathrm{Tr}[SF - 2\hat{\tau}(q)F^2 + F] = 0$

and hence the result. ■

APPENDIX C: ON ASYMPTOTICS IN THE TAILS

From Lemma 2 we have

$$\hat{\tau}(q) = \frac{\mathrm{Tr}[(S+I)F]}{2\,\mathrm{Tr}[F^2]},$$

and writing $F = q(q^{-1}A_1 - A_2)$, then

$$\lim_{|q|\to\infty} \hat{\tau}(q)q = \frac{\mathrm{Tr}[(S+I)A_2]}{\mathrm{Tr}[A_2^2]} = \frac{\sum_i a_{i,i}\,\mu_i^2 + 2\sum_{i\neq j}\sum a_{i,j}\,\mu_i\,\mu_j + \sum_i a_{i,i}}{\sum_i a_{i,i}^2 + 2\sum_{i\neq j}\sum a_{i,j}^2}, \tag{C.1}$$

where the $a_{i,j}$ and μ_i are the elements of A_2 and μ. Because the $a_{i,j}$ and μ_i are finite we write

$$b_1 = \frac{\sum_i a_{i,i}\,\mu_i^2}{N}; \qquad b_2 = \frac{\sum_{i\neq j}\sum a_{i,j}\,\mu_i\,\mu_j}{\frac{1}{2}N(N-1)}; \qquad b_3 = \frac{\sum_i a_{i,i}}{N};$$

$$b_4 = \frac{\sum_i a_{i,i}^2}{N} \quad \text{and} \quad b_5 = \frac{\sum_{i\neq j}\sum a_{i,j}^2}{\frac{1}{2}N(N-1)},$$

and so

$$\hat{\tau}(q)q = \frac{Nb_1 + N(N-1)b_2 + Nb_3}{Nb_4 + N(N-1)b_5},$$

and

$$\lim_{|q|\to\infty} \hat{\tau}(q)q = O(1).$$

The relevant terms in $Q^{(j)}(\hat{\tau})$ and $P^{(k)}(\hat{\tau})$ are

$$\hat{G}^{-1}F = \hat{\tau}[q^{-1}\hat{\tau}I - 2(q^{-1}A_1 - A_2)]^{-1}(q^{-1}A_1 - A_2),$$
$$\hat{G}^{-1}S = \hat{\tau}q^{-1}[q^{-1}\hat{\tau}I - 2(q^{-1}A_1 - A_2)]^{-1}S,$$
$$\hat{G}^{-1}A_2 = \hat{\tau}q^{-1}[q^{-1}\hat{\tau}I - 2(q^{-1}A_1 - A_2)]^{-1}A_2 \tag{C.2}$$

and substitute back into their definitions. Consequently, noting extraneous $\hat{\tau}$ and q^{-1} terms cancel, the asymptotic nature is determined by examining

$$\lim_{|q|\to\infty}[q^{-1}\hat{\tau}(q)I - 2(q^{-1}A_1 - A_2)]^{-1}(q^{-1}A_1 - A_2) = [q^{-1}\hat{\tau}(q)A_2^{-1} - 2I]^{-1},$$
$$\lim_{|q|\to\infty}[q^{-1}\hat{\tau}(q)I - 2(q^{-1}A_1 - A_2)]^{-1}S = [q^{-1}\hat{\tau}(q)S^{-1} - 2A_2S^{-1}]^{-1},$$
$$\lim_{|q|\to\infty}[q^{-1}\hat{\tau}(q)I - 2(q^{-1}A_1 - A_2)]^{-1}A_2 = [q^{-1}\hat{\tau}(q)A_2 - 2I]^{-1}.$$

Noting $\hat{\tau}(q)q = O(1)$ we thus have

$$\lim_{|q|\to\infty}\mathrm{Tr}[(\hat{\tau}(q)q\hat{G}^{-1}F)^j(\hat{\tau}(q)q\hat{G}^{-1}A_2)(j\hat{\tau}(q)q\hat{G}^{-1}S + I)] = O(N);$$
$$\lim_{|q|\to\infty}\mathrm{Tr}[(\hat{\tau}(q)q\hat{G}^{-1}F)^k(j\hat{\tau}(q)q\hat{G}^{-1}S + I)] = O(N), \tag{C.3}$$

that is, we have shown

$$\lim_{|q|\to\infty}\begin{Bmatrix} Q^{(j)}(\hat{\tau}(q)) \\ P^{(k)}(\hat{\tau}(q)) \end{Bmatrix} = O(N) \quad \forall j, k,$$

as required. ∎