A Selective Survey of Peter's Contributions to Finite Sample Theory

John C. Chao[1] and Grant H. Hillier[2]

[1]University of Maryland, [2]University of Southampton

I. Bibliography

A. Papers on Exact Distribution Theory

1. “The Exact Finite Sample Density of Instrumental Variable Estimators in an Equation with $n + 1$ Endogenous Variables," *Econometrica*, Vol. 48, No.4, May 1980. pp. 861-878.
2. “Exact Small Sample Theory in the Simultaneous Equations Model," Chapter 8 and pp. 449-516 in M.D.Intriligator and Z. Griliches (eds.) *Handbook of Econometrics*, North-Holland, 1983.
3. “The Exact Distribution of LIML:I," *International Economic Review*, Vol. 25, No. 1, February 1984, pp. 249-261.
4. “The Exact Distribution of the Stein Rule Estimator," *Journal of Econometrics*, Vol 25, No. 1/2, May/June, 1984, pp. 123-131.
5. “The Exact Distribution of Exogenous Variable Coefficient Estimators," *Journal of Econometrics*, Vol. 26, No. 3, December 1984, pp. 387-398.

6. “The Distribution of Matrix Quotients," *Journal of Multivariate Analysis*, Vol. 16, No.1, February 1985, pp. 157-161.

7. “The Exact Distribution of LIML:II," *International Economic Review*, Vol. 26, No. 1, February 1985, pp. 21-36.

8. “The Exact Distribution of the SUR Estimator," *Econometrica*, Vol. 53, No. 4, July 1985, pp. 745-756.

9. “Fractional Matrix Calculus and the Distribution of Multivariate Tests," in I.B. MacNeill and G.J. Umphrey (eds.), *Time Series and Econometric Modeling*, Dordrecht: D. Reidel, 1986, pp. 219-234.

10. “The Distribution of FIML in the Leading Case," *International Economic Review*, Vol. 27, No. 1, February 1986, pp. 239-243.

11. “The Exact Distribution of the Wald Statistic," *Econometrica*, Vol. 54, No. 4, July 1986, pp. 881-895.

12. “An Everywhere Convergent Series Representation of Hotelling’s Generalized *T*," *Journal of Multivariate Analysis*, Vol. 21, No.2, April 1987, pp. 238-248.

13. “Spherical Matrix Distributions and Cauchy Quotients," *Statistics and Probability Letters*, Vol. 8, 1989, pp. 51-53.

14. "Operational Algebra and Regression *t*-Tests" in P.C.B.Phillips (Ed.) *Models, Methods and Applications of Econometrics: Essays in Honor of A.R.Bergstrom*. Oxford: Basil Blackwell, 1993.

15. "Some Exact Distribution Theory for Maximum Likelihood Estimators of Cointegration Coefficients in Error Correction Models," *Econometrica*, Vol. 62, No. 1, January 1994, pp. 73-94.

16. "Posterior Distributions in Limited Information Analysis of the Simultaneous Equations Model Using the Jeffreys Prior" (with John C. Chao), *Journal of Econometrics*, Vol 87, November, 1998, pp. 49-86.

17. "Jeffreys Prior Analysis of the Simultaneous Equations Model in the Case with $n + 1$ Endogenous Variables" (with John C. Chao), *Journal of Econometrics*, Vol 111, No. 2, December, 2002, pp. 251-283.

18. "A Remark on Bimodality and Weak Instrumentation in Structural Equation Estimation," *Econometric Theory*, Vol. 22, No. 5, 2006, pp. 947-960.

19. "Exact Distribution Theory in Structural Estimation with an Identity," *Econometric Theory*, 2007 (forthcoming).

B. Papers on Partial and Weak Identification

20. “Partially Identified Econometric Models," *Econometric Theory*, Vol. 5, No. 2, August 1989, pp. 181-240.
21. “Asymptotic and Finite Sample Distribution Theory for IV Estimators and Tests in Partially Identified Structural Equations" (with In Choi), *Journal of Econometrics*, Vol 51, No. 1/2, January/February, 1992, pp. 113-150.
22. “GMM with Many Moment Conditions," (with Chirok Han), *Econometrica*, Vol. 74, No. 1, January 2006, pp. 147-192.

C. Papers on Asymptotic Expansions and Various Approximation Methods

23. "Approximations to Some Finite Sample Distributions Associated with a First Order Stochastic Difference Equation," *Econometrica*, Vol. 45, No. 2, March 1977, pp. 463-485.

24. "A General Theorem in the Theory of Asymptotic Expansions as Approximations to Finite Sample Distributions of Econometric Estimators," *Econometrica*, Vol. 45, No. 6, September 1977, pp. 1517-1534.

25. "An Approximation to the Finite Sample Distribution of Zellner's Seemingly Unrelated Regression Estimator," *Journal of Econometrics*, Vol 6, No. 2, September 1977, pp. 147-164.

26. "Edgeworth and Saddlepoint Approximations in a First Order Non-Circular Autoregression," *Biometrika*, Vol. 65, No. 1, February 1978, pp. 91-98.

27. "A Saddlepoint Approximations to the Distribution of the k-Class Estimator of a Coefficient in a Simultaneous System," (with A. Holly), *Econometrica*, Vol. 47, No. 6, November 1979, pp. 1527-1548.

28. “Finite Sample Theory and the Distributions of Alternative Estimators of the Marginal Propensity to Consume," *Review of Economic Studies*, Vol. 47, No. 1, January 1980, pp. 183-224.

29. “Best Uniform and Modified Padé Approximations of Probability Densities in Econometrics," Chapter 5 in W. Hildenbrand (ed.), *Advances in Econometrics*, Cambridge University Press, 1982, pp. 123-167.

30. “Marginal Densities of Instrumental Variable Estimators in the General Single Equation Case," *Advances in Econometrics*, Vol. 2, 1983, pp. 1-24.

31. “ERA’s: A New Approach to Small Sample Theory," *Econometrica*, Vol. 51, No. 5, September 1983, pp. 1505-1525.

32. “Finite Sample Econometrics Using ERA’s," *Journal of Japan Statistical Society*, Vol. 14, No. 2, November 1984, pp. 107-124.

33. “Large Deviation Expansions in Econometrics," in D. Slottje (ed.), *Advances in Econometrics*, Vol. 5, 1986, pp. 199-226.

34. “Asymptotic Expansion for Non-stationary Vector Autoregressions," *Econometric Theory*, Vol. 3, No. 1, April 1987, pp. 45-68.

35. “On the Formulation of Wald Tests of Non-Linear Restrictions," (with Joon Y. Park) *Econometrica*, Vol. 56, No. 5, September 1988, pp. 1065-1084.

36. “Higher Order Approximations for Frequency Domain Time Series Regression," (with Zhijie Xiao) *Journal of Econometrics*, Vol 86, No. 2, October 1998, pp. 297-336.

37. “Higher Order Approximations for Wald Statistics in Time Series Regressions with Integrated Processes," (with Zhijie Xiao) *Journal of Econometrics*, Vol 108, No. 1, May 2002, pp. 157-198.

38. “Second Order Expansions for the Distribution of the Maximum Likelihood Estimator of the Fractional Difference Parameter," (with Offer Liberman), *Econometric Theory*, Vol. 20, No. 3, 2004, pp. 464-484.

39. “Expansions for Approximate Maximum Likelihood Estimators of the Fractional Difference Parameter," (with Offer Liberman), *Econometrics Journal*, Vol. 8, No. 3, December 2005, pp. 367-379.

40. “Refined Inference on Long Memory in Realized Volatility," (with Offer Lieberman), *Econometrics Reviews*, 2007 (forthcoming).

41. “A Complete Asymptotic Series for the Autocovariance Function of a Long Memory Process," (with Offer Lieberman), *Journal of Econometrics* (forthcoming).

II. Exact Distribution Theory: An Overview

A. Some General Comments

1. Unlike asymptotic distribution theory, exact distribution theory does not come equipped with a toolbox of theorems that have very general applicability. Prior to Peter's work in this area, there was little general methodology, and each problem had to be tackled afresh with methods of attack that were devised 'on the job'.
2. What Peter has done starting with his seminal 1980 *Econometrica* paper is to show that it is possible to come up with a more general technical machinery for resolving exact distributional problems and to obtain general results. Indeed, Peter's work on operator algebra/fractional matrix calculus and on Padé and other rational approximants have given us some of the very few general methods for handling these types of problems.

B. **Exact Density of the IV Estimator: Phillips (1980)**

1. Obtained the exact density of IV/2SLS estimator in the completely general case, encompassing all previously obtained results.
2. Made use of powerful methods from multivariate analysis involving invariant polynomials of matrix arguments which were new at the time and which were little known even to mathematical statisticians, let alone econometricians.
3. A major innovation is the clever use of a matrix differential operator representation to extract the exact density in the final step.
4. The exact distribution of the IV estimator is shown to have moments that exist up to the degree of overidentification.
5. Derived an approximate form for the exact pdf based on a Laplace-type approximation which was also new at the time. This approximation gave a remarkably simple and computable form of the pdf, so that the effects of key parameters like the degree of overidentification could be easily calculated. Another nice feature of this approximation is that it preserves the tail behaviour of the exact density.

C. **Handbook of Econometrics Monograph on Small Sample Theory**: **Phillips (1983)**

1. This chapter became the standard reference for young econometricians who want to learn about finite sample theory for the SEM.

2. Unusual for that period, this chapter actually contains sections on hypothesis testing and confidence sets.

D. **Exact Distributions of LIML and FIML**: **Phillips** (**1984**, **1985**, **1986**)

1. The new methods of multivariate analysis used for the IV estimator are also applied to LIML to derive results in complete generality.
2. Derivation of the exact distribution of LIML is harder because the estimator is only implicitly defined (not an explicit function of the underlying sufficient statistics).
3. The exact distribution of LIML is found to have Cauchy-like tails, so that no integer moment exists.
4. An interesting point in these papers is the beginning of a discussion of the role of identification in determining both the exact and asymptotic properties of estimators. In particular, when $\Pi_2 = 0$, the so-called leading case where β is completely unidentified, it was observed in Phillips (1984) that the exact densities of LIML, LIMLK, and 2SLS do not involve β and are all invariant to the sample size, reflecting the fact that the estimators in this case carry no information (however defined) on the parameter of interest.

Exact Distributions of LIML and FIML (cont'd)

5. A technical advance put forth in Phillips (1986) is the realization that it may be possible to deduce the density of an estimator from the invariance properties of the underlying statistics upon which it depends, together with those of the objective function. This has led to alternative derivations for the exact densities of both LIML and FIML in the leading case that are greatly simplified and more intuitive.

E. **Exact Distribution of the Wald Statistic: Phillips (1986)**

1. Phillips (1986) derives the exact distribution of the Wald statistic for testing general linear restrictions on coefficients of a multivariate linear model.

2. A general formula for the exact density was provided which encompasses both the null distribution of the statistic and the non-null distribution obtained under a generally specified alternative hypothesis.

3. The result obtained subsumes all previous results on the Wald statistic and contains as special cases results for the F statistic, the Hotelling T^2 statistic, and the Hotelling generalized T_0^2 statistic.

Exact Distribution of the Wald Statistic (cont'd)

4. A key part of the analysis leading to the non-null distribution requires the derivation of the conditional density of

$$w = N(Dvec(A^*) - d)'\{D(C \otimes M)D'\}^{-1}(Dvec(A^*) - d)$$

given $C = Y'(I - P_X)Y$. This is a non-central quadratic form in normal variates - certainly, an extremely complicated statistic for which to derive the exact distribution! Ruben (1962) had earlier obtained a recursively defined (not explicit) expression for the density of such a quadratic form (with a fixed matrix), but Phillips (1986) was the first to derive an explicit form for this density.

5. The unconditional density of the Wald statistic w is then obtained by averaging the conditional distribution with respect to the density of C. This is a very difficult integration problem, and Peter was able to evaluate this integral using the fractional matrix operator methods that he had developed. Operational calculus had been used in applied mathematics, for example, to solve differential equations. This paper, along with Peter's 1980 IV paper and the papers on SUR and the Stein estimator, generalized operational calculus to the matrix case and pioneered its use in statistics and econometrics.

F. **Bayesian Analysis under the Jeffreys Prior: Chao and Phillips (1998, 2002)**

1. A strong correspondence was found between the exact finite sample distribution of LIML and the marginal posterior distribution of β under the Jeffreys prior.

2. In the case of a canonical model under just identification, the posterior pdf has exactly the same functional form as the exact density of LIML.

3. In the overidentified case, the posterior distribution under the Jeffreys prior still has the same Cauchy-like tails as the sampling density of LIML.

III. Partial Identification

- **Papers: Phillips (1989) and Choi and Phillips (1992)**
- Peter’s 1989 paper can legitimately be viewed as marking the beginning of the resurgence of interest in the SEM and its attendant identification issues, although many contributors to the recent literature on “weak identification" have failed to acknowledge this.
- Spelled out in detail how lack of identification impacts on standard asymptotics. In particular, both point estimators and test statistics in this case have asymptotic distributions that are either identical, or are closely related, to the finite sample distributions that obtain under normality.
- The partial identification results, thus, point to the value of finite sample analysis. More specifically, finite sample analysis not only gives us distributional results that hold exactly under certain strong assumptions (usually Gaussian underlying distributions) but in situations where the primitives of the problem satisfy some CLT, this type of analysis can also be quite informative about the behaviour of statistics in large sample and under non-normality of the underlying distributions.

- Phillips (1989) and Choi and Phillips (1992) are really the first to study the behaviour of test statistics under identification failure. They show that in this case Wald statistics are no longer chi-square distributed under the null and indeed have the same asymptotic distribution under both the null and the alternative and are, thus, inconsistent.

- Perhaps a more subtle problem with the Wald test, as Dufour (1997) and others have noted, is that, for many models that are *locally almost identified* (in the terminology of Dufour, 1997), confidence sets obtained by inverting the Wald statistic will have confidence level $= 0$ (regardless of the nominal level). For the partially identified SEM, this result can be shown to be a direct implication of results obtained in Phillips (1989).

Example

- Consider the case where β is a scalar parameter, and let

$$W_T(\beta_a) = \frac{\left(\widehat{\beta}_{IV,T} - \beta_a\right)^2}{\widehat{\sigma}^2 [y_2'(P_Z - P_{Z_1})y_2]^{-1}}$$

be the Wald statistic for testing $H_0 : \beta = \beta_a$. In this case, Theorem 2.8 part (a) of Phillips (1989) specializes to

$$W_T(\beta_a) \Rightarrow \frac{W(r - \beta_a)^2}{1 + r^2} \quad \text{under } H_0,$$

where $W \sim \chi^2_{k_2}$ and r has a mixed normal distribution not depending on β_a. Now, a confidence set $C_{\beta,T}$ corresponds to the collection of β_a's such that $H_0 : \beta = \beta_a$ is not rejected when tested against $H_1 : \beta \neq \beta_a$.

Hence, if c_α is a constant such that $P(\chi_1^2 \le c_\alpha) = 1 - \alpha$, then,

$$\liminf_{T\to\infty} \inf_{\beta_a \in \mathrm{R}} P_{\beta_a}\{\beta_a \in C_{\beta,T}\} = \liminf_{T\to\infty} \inf_{\beta_a \in \mathrm{R}} P_{\beta_a}\{W_T(\beta_a) \le c_\alpha\}$$

$$\le \inf_{\beta_a \in \mathrm{R}} \lim_{T\to\infty} P_{\beta_a}\{W_T(\beta_a) \le c_\alpha\}$$

$$= \inf_{\beta_a \in \mathrm{R}} P_{\beta_a}\left\{\frac{W(r-\beta_a)^2}{1+r^2} \le c_\alpha\right\}$$

$$\le \lim_{|\beta_a|\to\infty} P_{\beta_a}\left\{\frac{W(r-\beta_a)^2}{1+r^2} \le c_\alpha\right\} = 0$$

since

$$\frac{W(r-\beta_a)^2}{1+r^2} = \frac{W\left(\frac{r^2}{\beta_a^2} - 2\frac{r}{\beta_a} + 1\right)\beta_a^2}{1+r^2} \overset{p}{\to} \infty \text{ as } n \to \infty.$$

Conclusion

1. Thank you for helping us better understand the finite sample behaviour of estimators and test statistics.

2. Thank you for providing us with new analytical tools for conducting our research.

3. Happy 60th Birthday!!!