Econometric Theory, **19**, 2003, 495–511. Printed in the United States of America.
DOI: 10.1017/S0266466603193085

VISION AND INFLUENCE IN ECONOMETRICS: JOHN DENIS SARGAN

Peter C.B. Phillips
Cowles Foundation, Yale University
University of Auckland
University of York

Denis Sargan's intellectual influence in econometrics is discussed and some of his visions for the future of econometrics are considered in this memorial article. One of Sargan's favorite topics in econometric theory was finite sample theory, including both exact theory and various types of asymptotic expansions. We provide some summary discussion of asymptotic expansions of the type that Sargan developed in this field and give explicit representations of Sargan's formula for the Edgeworth expansion in the case of an econometric estimator that can be written as a smooth function of sample moments whose distributions themselves have Edgeworth expansions.

A student is like green grass and a great teacher is like the spring sun. The benefit from the sun is infinite, and little grass can hardly pay it back, although it tries its best.

—Chinese saying

1. LESSONS IN RESEARCH

In an era where the half-life of academic research can often be measured in a matter of months, Denis Sargan's published papers show a remarkable durability. A quick dip into a Sargan paper reveals concerns (such as dynamic specification, the marriage of simultaneity and serial dependence, or the finite sample properties of econometric estimators) that are as relevant today as they were at the time of writing. Sargan's papers are filled with technical innovations, and much of his work shows little sign of aging even after decades of subsequent research. In this sense, Sargan's research contributions and the concerns that motivated his work have become classics of econometric literature.

Parts of Section 1 were presented in March 2002 in the author's Sargan Lecture at the Royal Economics Society Conference, Warwick, UK. My thanks go to John Chao, David Hendry, Essie Maasoumi, Peter Robinson, and Katsumi Shimotsu for helpful comments on the original version of this paper. I learned the Chinese saying that heads this article from Sainan Jin. Thanks also go to the NSF for research support under grant SES 00-92509.
Address correspondence to: Peter C.B. Phillips, Yale University/Cowles Foundation, 30 Hillhouse Ave., New Haven, CT 06511, USA; e-mail: peter.phillips@yale.edu.

Most of Sargan's major works were published in *Econometrica* and were therefore widely accessible when published and are now conveniently accessible online. Nonetheless, much of Sargan's thinking about econometrics was never published. Fortunately, because of his extraordinary pedagogical impact over several decades at the London School of Economics (LSE) and more generally within the econometrics community in the United Kingdom, many of his ideas and techniques have become part of the intellectual fabric of econometrics. Naturally enough, Sargan's intellectual legacy remains strongest in the UK, although his former students are now widely scattered internationally.

Sargan's students and colleagues at the LSE had the added advantage of watching his mind work at close quarters—in seminars and in valuable dyadic conversations with him. In my own case, I came from New Zealand to do research at the LSE and knew many of his published contributions well. But it was a surprise for me to discover the rich vein of material in his unpublished papers. My first short meeting with Denis took place in his LSE office in October 1971. In what was a characteristically unassuming way, Denis opened a cabinet and revealed a dusty bottom drawer that was full of unpublished papers. He pulled out a selection of these, saying, "These will give you an idea of the sort of thing we are working on at LSE."

To the uninitiated, a Sargan working paper often looked like an impenetrable jungle of mathematics and conceptual exposition. Formulas ran for pages, algebraic symbols carried strange decorations, notation might be assumed and undefined, unusual branches of mathematics might be called upon without remark, theorems might be cited without reference as if they were as familiar to the reader as to the author, and the mathematics might be entangled in a high level discussion that confronted some deep conceptual issues of econometric modeling. These papers required a sustained effort to read and an enormous commitment to master that involved months of devoted line-by-line reading. Looking back now on my own experience with them, the effort proved to be a valuable and long-lived investment.

All of Sargan's working papers were at the cutting technical edge of the subject. They dealt with difficult topics like full information maximum likelihood (FIML), Edgeworth expansions, and stochastic differential equations, all in an imaginative way that was very different from what I had seen in the literature up to 1970. By the time I had finished reading these working papers, a lot of new econometrics had come my way. More significantly, I had learned some memorable new lessons about research:

(i) *No Problem Is Too Difficult to Tackle*. Sargan's papers revealed that no problem was too difficult to tackle. With the right approach, conceptualization, and mathematics you can always make headway even when things may at first seem impossible. Throughout his career, Sargan continued to surprise the econometrics community by tackling problems that seemed impossible, looking at problems that other researchers would hardly contemplate as tractable, and often doing so with new and astonishing results. For example, whereas other researchers were doing what was eminently possible by worrying about first-order conditions and a precise limit theory for FIML using central limit theory and nonlinear regression, Sargan was doing the apparently impossible by developing a finite sample theory and refined asymptotics using Edgeworth expansions. This powerful example inspired all of us who stood around him in the UK.

(ii) *Problems Are Often Simplified by Generalization*. By formulating a problem in a general way and developing the appropriate mathematics to handle that generality, it is often possible to achieve substantial simplifications. This is a principle we encounter every day when we are teaching elegant methods that have been developed in the past, the projection matrix algebra of the general linear model that was worked out by Aitken (1935) being one prominent example in econometrics. In spite of such constant reminders, one sees a widespread reluctance to implement this principle in research. Of course, it requires imagination and courage to make the conceptual and mathematical leaps that are involved in successful generalization, and it is always safer to follow established paths. Yet, time and again throughout his career, Sargan demonstrated the power of the principle of generalization, many of his papers making highly imaginative technical and conceptual leaps forward.

One fascinating example was his finite sample theory for the FIML estimates of the structural coefficients and demonstration of the fact that these estimates have Cauchy-like tails, a property that we now know extends to cointegration vector estimates that are obtained by reduced rank regression (Phillips, 1994). Another example was Sargan's remarkable three-page proof that the FIML estimates of the reduced form coefficients have finite sample moments to order $T - m - n$ (i.e., the sample size less the number of variables in the system), whereas those obtained by other methods (such as instrumental variables) usually have no moments and heavier tails with greater probabilities of outliers, a fact that is of importance in forecasting from structural systems. These results, among many others, were presented at study group and workshop meetings at the LSE in the early 1970s. The structural form FIML paper was first given at the 1970 World Congress of the Econometric Society. Versions of the results eventually appeared in Sargan (1988b, Chs. 3 and 6).

Sargan usually developed the mathematics he needed as he went along, so that his articles became voyages of discovery in technique in addition to purveyors of econometric methodology and applications. He was particularly creative with matrix algebra, one example being his constructive use of 0,1 matrices like selectors, duplicators, commutators, and permutators, long before they were systematically studied in the mathematical and statistical literatures.

(iii) *Don't Shirk Algebraic Complexity*. Sargan's papers (particularly his unpublished papers) reveal a love of algebraic complexity. Some researchers find it convenient to gloss over algebraic detail as tiresome. This practice is encouraged by modern symbolic computation and is often endorsed by journal editors who are working under space limitations and who want to achieve page reductions in lengthy papers. In contrast, Sargan seemed to delight in filling pages of a paper with long strings of algebraic calculation, sometimes going well over a page, devising intriguing new notation as he went along. One senses that it was the intimate familiarity with these algebraic details, even in ferociously complicated cases, that enabled Sargan to make headway on difficult problems. Grinding your way through a lot of algebraic detail to reach completion may appear like a very unimaginative exercise, but sometimes it is this very process that enriches the understanding of a problem to the stage where major conceptual leaps become possible. In what sense then is it better to pretend that the details and the hard work don't exist?

(iv) *Maintain an Active Concern for Economic Issues*. Econometricians do well to keep an active interest in economic issues and a concern for statistical methods that are suitable for economic models and data. Interestingly, Sargan never seemed to get preoccupied with probability or statistics. One never got the impression that Sargan was writing a paper for statisticians or probabilists. He always seemed to be writing papers on econometric methods for use in economics, maintaining a focus on econometric problems and the estimation of economic relations. In fact, his papers often have these very words in their titles. In recent years, this clarity of focus is much less common. The interface between econometrics and statistical theory is now much more fluid than it was in the 1950s, 1960s, and 1970s, when Sargan did his major work. Econometricians now do research that is often identified as easily, if not more easily, with the subject of statistics than it is with economics. Many publish their work in probability and statistics journals. It is, of course, important that econometric technique, methodology, and proof be evaluated in the wider circle of statistical research, a principle that was well understood by prominent early researchers like Francis Edgeworth, Irving Fisher, and Tjalling Koopmans. But what is ultimately more important is that the connection with economics be maintained as a strong one, something these early researchers never forgot. Sargan's research also did just that, by concentrating on issues that he felt to be of importance to economic researchers and by keeping an eye on empirical economic applications.

2. SARGAN'S VISION FOR THE FUTURE OF ECONOMETRICS

Sargan sought not just to do econometrics and do it very well; he also had a vision of where it could most fruitfully go in the future. Three areas where that vision was most powerful were in dynamic specification, identification, and improving inference. These topics manifested themselves repeatedly in Sargan's major contributions over the period 1960–1985.

2.1. Better Dynamic Specification

Sargan's (1964) Colston Conference paper was largely motivated by empirical issues of dynamic specification that arose in an applied study of wage inflation in the UK. In such applications, as had earlier been emphasized in the work of Sargan's LSE colleague Bill Phillips (1956, 1958), the dynamic time form of responses plays a crucial role in the economic impact of phenomena like inflation and, in consequence, policy responses to it. Sargan's contribution was seminal for many reasons, which are examined in detail by David Hendry elsewhere in this memorial issue and which we only mention in passing here. What has proved perhaps to be the most significant contribution from the econometric perspective is that, in seeking to provide a foundation for empirical specification searches, Sargan laid out the elements of a practical methodology for specifying dynamic relations that involves working from general forms to more specific formulations and sought to provide a justification for the econometric testing that plays a role in this reduction process. That mantle of methodological research was taken up in earnest at the LSE in the 1970s and 1980s and has, in consequence, become known as the "LSE approach," under the leadership of David Hendry in an extensive agenda of applied and theoretical econometric research.

These themes continued to be present in Sargan's later research, including his efforts on testing for the presence of common factors in the equation errors and dynamics (Sargan, 1977b, 1980a). We now commonly think of these procedures within the wider context of econometric model selection, a literature that arose from the early work of Akaike (1973), Leamer (1978), Schwarz (1978), and Rissanen (1978) and that connects in important ways to Bayesian methods of model choice where there are even earlier antecedents (e.g., Jeffreys, 1961). Indeed, one of Sargan's later long-unpublished works (2001) on significance testing takes account of these contributions and, in particular, the issue of consistent model choice (or completely consistent significance testing procedures) in developing rules for increasingly conservative statistical tests whereby the critical values expand at logarithmic rates with the sample size, so that the probability of type I and type II errors goes to zero asymptotically.

Dynamic specification now forms part of the core subject matter of time series econometrics. The issues arising from it have become no less important over time, although they may take on different forms, as the subject has moved on to include nonparametric and semiparametric methods and to encompass the nonstationarity and long memory properties of economic data. Although there is no explicit mention of it in Sargan's work, the data-driven approaches to modeling and specification searching (e.g., Andrews, 1991; Hendry, 2001; Hendry and Krolzig, 1999, 2001, 2002; Phillips, 1992, 1995a, 1995b, 1996, 1999, 2003; Robinson, 1991) that are now being implemented in practical work seem to be very much in the spirit of Sargan's tradition.

2.2. Recognizing the Effects of Near Lack of Identification

Early researchers on simultaneous equations methodology recognized the practical difficulties of assessing identification. Tests for underidentification, like the Koopmans and Hood (1953) test, were a manifestation of this concern. Later, Sargan (1958) developed a version of this test that was applicable with instrumental variable estimation, and generalized method of moments (GMM) versions are also possible (Arellano, Hansen, and Sentana, 1999). In practical work, however, these tests are seldom used, and empirical research generally proceeds as long as an equation is apparently "identified" by order conditions or some other restrictions. Sargan recognized that such practices were potentially hazardous and that, in the event of near lack of identification, the asymptotic properties of econometric estimators and tests would be affected. In a paper given to a Royal Statistical Society study group in 1975 that was eventually published in Sargan (1988a, Ch. 12), he explored the relationship between identification and consistent estimability in systems of simultaneous stochastic equations. Later, in his presidential address to the Econometric Society in 1980, Sargan (1983a) looked at nonlinear in parameter models that were nearly unidentified in the sense that the first-order rank condition for local identification failed but higher order shape conditions held so that there was still identification. In 'singular' cases like these, Sargan found that the conventional asymptotic theory for instrumental variables estimation broke down, with lower rates of convergence and a nonnormal limit theory applying. A subsequent paper, Sargan (1983b), showed that similar problems of singularity occurred in dynamic models with autoregressive errors. Although a general limiting distribution theory was not given in that case, Sargan remarked that "in finite samples the distributions of estimators derived from models which are almost singular tend to approximate those from models which are exactly singular."

This work by Sargan on near lack of identification anticipated much future research. A limit theory for instrumental variables estimates in exactly singular (or totally unidentified) systems was given a few years later in Phillips (1989). In this limit theory, instrumental variable estimates converge weakly to random variables, reflecting in the limit the uncertainty about the coefficients that is implicit in their lack of identification. This new limit theory is similar to what has been found to apply in the weakly identified cases that have been considered more recently (e.g., Staiger and Stock, 1997). As Sargan (1983b) put it, when there is a failure of the first-order conditions for local identifiability, "the resulting estimates of the model will not possess the usual asymptotic normality properties." There are also major effects on statistical testing, and some procedures are now known to better reflect the uncertainty that is implied by lack of identification or near lack of identification than others (Kleibergen, 2000; Forchini and Hillier, 2002; Moreira, 2001). We further know that increasing the number of instruments in a controlled way can help to produce consistent estimates, thereby compensating for the fact that the instruments are individually weak, or in some cases, even irrelevant (Chao and Swanson, 2002; Han and Phillips, 2002). As this field of research deepens and researchers become more aware of earlier thinking, Sargan's seminal contributions to the subject will hopefully become better recognized.

2.3. Improving Inference

The finite sample properties of econometric estimators and tests were a concern to researchers even in the early days of the Cowles Commission studies of simultaneous equations. New approaches to fitting equations were typically justified, as they are today, by asymptotic arguments, whereas the behavior of econometric methods in finite samples is what is most relevant in applications. Sargan realized that limited data and model complexity both strain conventional asymptotic justifications. In a research agenda that spanned nearly 25 years from the early 1960s, Sargan sought to find methods by which finite sample inference in econometrics might be improved. Using exact distribution theory, asymptotic expansions, moment approximations, and simulation methods he provided analyses of the finite sample properties of econometric estimators in a wide range of settings.

Sargan's Walras–Bowley lecture to the North American Meetings of the Econometric Society in 1974 (published in *Econometrica*, 1976) summarized much of his research on this topic and was distinctive because of the generality of the approach as well as the variety of the methods it suggested. Most obvious were the general formulas he gave for Edgeworth expansions of econometric estimators, which were based on an earlier algorithm developed by Chambers (1967). The constructive process underlying this algorithm is detailed in Section 4. Sargan's dream was that general formulas of this type could be incorporated into regression software and, with the deployment of estimated coefficients, used to adjust critical values and improve inference.

As it turned out, however, this dream was never realized, for it soon became clear that Edgeworth approximations were not accurate enough to be used in empirical research. Indeed, in many cases they do not improve on first-order asymptotic theory. Somewhat ironically, when the first-order asymptotic theory is itself poor, this failing of the Edgeworth approximation is particularly noticeable. In such cases the gap that needs to be made up by the Edgeworth approximation is too large for its higher order terms to perform satisfactorily. The problem was shown to be especially acute in dynamic models (Phillips, 1977), where the usual asymptotics progressively break down as the nonstationarity zone is approached. In consequence, Sargan's dream of practical implementation and widespread use of Edgeworth expansions was never fulfilled. Instead, asymptotic expansions of this type are now used mainly to explain the good and bad performance of first-order asymptotics and to justify simulation-based approximations like the bootstrap.

Sargan himself recognized the importance of simulation-based approaches to improving inference, suggesting in his Walras–Bowley lecture a parametric version of the bootstrap, which he called the Barnard approximation, following an earlier suggestion that was made by George Barnard. Sargan's development of this simulation-based approximation was distinctive because it made constructive use of first-order asymptotics through control variates to aid the simulation. Although none of these methods has since been systematically employed in applied work, the thrust of the research agenda to improve inference was clear, and his work marries well with recent developments on the bootstrap and its justification by way of asymptotic expansions, where the impact on applied research has been substantial.

3. MENTORSHIP

Doctoral students suffer a common plight as they search for thesis topics, agonize over technical problems, and seek guidance and inspiration from the literature and mentors. One of Sargan's major lifetime contributions was to help relieve this academic suffering. He had a deep and sincere desire to help people that was manifest in the help he gave all his Ph.D. students and younger colleagues, often writing out pages of mathematical derivations to help them formulate a problem and overcome technical obstacles. Sargan's generosity to his students and colleagues was legendary at the LSE and undoubtedly played a major role in attracting the large number of doctoral students in econometrics that he supervised. It is well illustrated by this personal story I heard from one of Sargan's students. Not being able to prove a result himself but knowing it was vital to his dissertation work, this student gave Sargan a "proof" that started out the argument, worked backward from the desired result, and filled in the middle with a load of rubbish. After a couple of weeks, Sargan called him in and gave him a new proof, saying simply, "I don't think the original one quite worked."

Stories like this one make it clear that in our small world of econometrics Sargan was an academic bodhisattva. According to Buddhist teaching, a bodhisattva is someone who gives up nirvana to relieve the suffering of others and help them on the path to enlightenment. Sargan most certainly performed both of these roles in good measure. Beyond this, Sargan's eminent good taste in research, his capacity for right thinking, and his principled leadership had an enormous indirect impact, instilling in all those around him the aspiration to go out and help others in their turn. This bodhicitta has produced a large progeny of grandstudent descendants who are now working econometricians in institutions all around the globe—an invisible college of econometricians that makes for an astonishing legacy of intellectual influence and practice.

4. CONSTRUCTING EDGEWORTH EXPANSIONS

What follows is a skeleton of the constructive process by which Edgeworth expansions are derived in a general, smooth case, explicating the algebra given in Sargan's (1976) original treatment that was itself based on Chambers (1967). The history of this construction is old, dating back to Edgeworth's original development (which even included functions of multivariate statistics, although this seems to have gone largely unnoticed), Wallace (1958), Phillips (1977, 1982), Withers (1983), and Hall (1992) in various notations and at various levels of generality.

Part of the difficulty of Edgeworth expansions is notational. The complexity of the formulas in general cases is formidable, strains intuition, and often leads to repeated algebraic errors, the latter problem affecting many published articles, including some of the final expressions given in Sargan (1976, 1977a). One advantage of the notation used in the following development is that the essential ingredients remain explicit in the final formulas, so that one can see the impact of the form of the econometric estimator and the distributional properties of the underlying data on the finite sample distribution.

Once the algebra is laid out in a suitably general form, it can all be done flawlessly by symbolic computer manipulation. This type of implementation was envisaged by both Chambers (1967) and Sargan (1976), but the programs available at the time were of limited capability. Symbolic manipulation has now advanced to the stage where quite general implementations of this type are possible. A recent book by Andrews and Stafford (2000) shows how this can be accomplished, developing Edgeworth expansions, Bartlett corrections, and Cornish–Fisher expansions at some level of generality.

Let $\theta_n(m)$ be a statistic (such as the error in an econometric estimator or a t-ratio type test statistic) with an asymptotic normal distribution. This statistic is dependent on m, a vector of sample moments with density $\mathrm{pdf}(m)$, and the sample size n. In the simplest cases, the function $\theta_n(m)$ will be known explicitly from the form of the estimator or the test statistic. When θ_n is an extremum estimator, the functional dependence $\theta_n(m)$ is obtained by inverting the power series development of the first-order conditions of the extremum problem, at least up to an appropriate order of approximation that will depend on how far we want to develop the Edgeworth series. It is similarly convenient to assume that m is appropriately centered about corresponding population moments so that $E(m) = 0$ and that $\theta_n(0) = 0$. We further assume that m has moments of a high enough order for the operations that follow to be valid, that the standardized moment $m^n = \sqrt{n}m$ has cumulants of the same order of magnitude in n as would a simple standardized mean of independent and identically distributed variates (i.e., the cumulants of nm are $O(n)$ as $n \to \infty$), and that $m^n \to_d N(0, V)$ with a nonsingular variance matrix $V = (v_{ij})$. The limit normal distribution of the standardized moment $m^n = \sqrt{n}m$ is the fountainhead from which the Edgeworth expansion for $\theta_n(m)$ springs.

In an attempt to make the final formulas easier to interpret we use the following notation:

$$\theta_n(m) := \sqrt{n}e(m) = \sqrt{n}e(m^n/\sqrt{n}),$$

$$\mathrm{cf}(x;\theta_n) := \int e^{ix\theta_n(m)}\mathrm{pdf}(m)\,dm = \text{characteristic function (cf) of } \theta_n,$$

$$\text{cf}(z; m^n) := \int e^{iz'm^n}\text{pdf}(m^n)\,dm^n = \text{cf of } m^n,$$

$$\text{cgf}(z; m^n) := \log[\text{cf}(z; m^n)] = \text{cumulant generating function (cgf) of } m^n,$$

$$\text{pdf}(r) := \frac{1}{2\pi}\int e^{-irx}\text{cf}(x; \theta_n)\,dx$$

$$= \text{probability density function (pdf) of } \theta_n,$$

$$e_r = \frac{\partial e(0)}{\partial m_r}, \qquad e_{rs} = \frac{\partial^2 e(0)}{\partial m_r \partial m_s}, \qquad \text{cgf}_{ab} = \frac{\partial^2 \text{cgf}(0; m^n)}{\partial z_a \partial z_b},$$

$$\text{cf}_{ab}(z; m^n) = \frac{\partial^2 \text{cf}(z; m^n)}{\partial z_a \partial z_b}, \text{ etc.}$$

We also employ the convenient notational summation device that a repeated subscript is summed over itself, e.g., $e_{rs}m_r m_s = \sum_r \sum_s e_{rs} m_r m_s$, and we assume that e has continuous derivatives to the order required by the necessary expansions. In first-order asymptotics, $\theta_n(m) \to_d N(0, \omega^2)$ with $\omega^2 = e^{0\prime}Ve^0 = v_{ij}e_i e_j$, where $e^0 = \partial e(0)/\partial m = (e_r)$ and where we assume $\omega^2 > 0$.

The constructive process for the Edgeworth expansion begins with the Taylor representation

$$\theta_n(m) = \sqrt{n}e(m) = \sqrt{n}\left\{e_r m_r + \frac{1}{2}e_{rs}m_r m_s + \frac{1}{6}e_{rst}m_r m_s m_t + O_p(n^{-2})\right\}$$

$$= e_r m_r^n + \frac{1}{2\sqrt{n}}e_{rs}m_r^n m_s^n + \frac{1}{6n}e_{rst}m_r^n m_s^n m_t^n + O_p(n^{-3/2}), \tag{1}$$

which is a polynomial in the components m_r^n with terms involving higher order powers of $n^{-1/2}$ that are of decreasing importance as $n \to \infty$. This polynomial representation produces a corresponding expansion for the characteristic function of $\theta_n(m)$. Proceeding in a purely formal operational way that is attentive to orders of magnitude, we get

$$\text{cf}(x; \theta_n) = \int e^{ix\theta_n(m^n)}\text{pdf}(m^n)\,dm^n$$

$$= \int e^{ixe_r m_r^n + (ix/2\sqrt{n})e_{rs}m_r^n m_s^n + (ix/6n)e_{rst}m_r^n m_s^n m_t^n + O(n^{-3/2})}\text{pdf}(m^n)\,dm^n$$

$$= \int e^{ixe_r m_r^n}\left[1 + \frac{ix}{2\sqrt{n}}e_{rs}m_r^n m_s^n + O(n^{-1})\right]\text{pdf}(m^n)\,dm^n,$$

and noting that

$$\mathrm{cf}_{ab}(z; m^n) := \int (i m_a^n)(i m_b^n) e^{iz' m^n} \mathrm{pdf}(m^n)\, dm^n,$$

we deduce that

$$\begin{aligned} \mathrm{cf}(x;\theta_n) &= \mathrm{cf}(xe^0; m^n) + \frac{ixe_{rs}}{2\sqrt{n}}\left(\frac{1}{i}\right)^2 \mathrm{cf}_{rs}(xe^0; m^n) + O(n^{-1}) \\ &= \mathrm{cf}(xe^0; m^n) - \frac{ixe_{rs}}{2\sqrt{n}}\,\mathrm{cf}_{rs}(xe^0; m^n) + O(n^{-1}). \end{aligned} \tag{2}$$

Formula (2) writes the characteristic function of θ_n as a series of linear forms involving the cf of m^n and its successive derivatives with coefficients depending on the derivatives of e. Note the presence of the complex constant $i = \sqrt{-1}$ in (2), which arises because the cf is, in general, complex.

The next step in the process is to expand the cf of m^n in a series about the cf of the limiting distribution of m^n. To achieve this, we first expand the cgf of m^n in the following Taylor series at the origin

$$\mathrm{cgf}(z; m^n) = \frac{1}{2}\,\mathrm{cgf}_{ij}\, z_i\, z_j + \frac{1}{6}\,\mathrm{cgf}_{ijk}\, z_i\, z_j\, z_k + O(z^4), \tag{3}$$

where we note that the first nonzero term is quadratic because m^n is centered about its mean value. By definition

$$\mathrm{cf}(z; m^n) = \exp\{\mathrm{cgf}(z; m^n)\}, \tag{4}$$

so that setting $z = xe^0$ in (3) and (4) we obtain

$$\begin{aligned} \mathrm{cf}(xe^0; m^n) &= \exp\left\{\frac{x^2}{2}\,\mathrm{cgf}_{ij}\, e_i\, e_j + \frac{x^3}{6}\,\mathrm{cgf}_{ijk}\, e_i\, e_j\, e_k + O(x^4)\right\} \\ &= \exp\left\{\frac{x^2}{2}\,\mathrm{cgf}_{ij}\, e_i\, e_j\right\}\left[1 + \frac{x^3}{6}\,\mathrm{cgf}_{ijk}\, e_i\, e_j\, e_k + O(x^4)\right]. \end{aligned} \tag{5}$$

Similar expressions now hold for the derivatives. In particular

$$\mathrm{cf}_{ab}(xe^0; m^n) = \exp\left\{\frac{x^2}{2}\,\mathrm{cgf}_{ij}\, e_i\, e_j\right\}[\mathrm{cgf}_{ab} + (\mathrm{cgf}_{aj}\, e_j)(\mathrm{cgf}_{bk}\, e_k)x^2 + O(n^{-1/2})], \tag{6}$$

where we use the fact that the cumulants of m^n of order p are $O(n^{1-p/2})$, so that $\mathrm{cgf}_{ijk} = O(n^{-1/2})$.

Using (5) and (6) in (2) we find the following series expansion for the cf of θ_n:

$$\begin{aligned}\mathrm{cf}(x;\theta_n) &= \exp\left\{\frac{x^2}{2}\,\mathrm{cgf}_{ij}e_ie_j\right\}\left[1-\frac{ix}{2\sqrt{n}}\,\mathrm{cgf}_{rs}e_{rs}\right.\\ &\quad\left.+\frac{x^3}{6}\left\{\mathrm{cgf}_{ijk}e_ie_je_k-\frac{3i}{\sqrt{n}}\,e_{rs}(\mathrm{cgf}_{rj}e_j)(\mathrm{cgf}_{sk}e_k)\right\}\right]+O(n^{-1})\\ &= \exp\left\{-\frac{x^2}{2}\,\omega^2\right\}\left[1-\frac{ix}{2\sqrt{n}}\,\mathrm{cgf}_{rs}e_{rs}\right.\\ &\quad\left.+\frac{x^3}{6}\left\{\mathrm{cgf}_{ijk}e_ie_je_k-\frac{3i}{\sqrt{n}}\,e_{rs}(\mathrm{cgf}_{rj}e_j)(\mathrm{cgf}_{sk}e_k)\right\}\right]+O(n^{-1})\end{aligned} \tag{7}$$

because $\omega^2 = v_{ij}e_ie_j = -\mathrm{cgf}_{ij}e_ie_j$, the variance of the limiting distribution of θ_n. Because (7) involves second and third derivatives of $\mathrm{cgf}(z;m^n)$ at the origin, some of the coefficients in (7) again involve the complex constant i.

Formula (7) is an asymptotic expansion of the cf of θ_n about the cf of its limit distribution, $\exp\{-\frac{1}{2}x^2\omega^2\}$. The approximation obtained by truncating the higher order terms of $O(n^{-1})$ is just a polynomial (in x) multiple of this asymptotic cf. By inverting (7) term by term we deduce the corresponding asymptotic expansion of the pdf of θ_n. To do so, we use the relation

$$(-1)^k\varphi^{(k)}(z) = \frac{1}{2\pi}\int (it)^k e^{-t^2/2}e^{-itz}\,dt$$

or, as is more appropriate here,

$$\frac{1}{\omega^{k+1}}H_k\left(\frac{r}{\omega}\right)\varphi\left(\frac{r}{\omega}\right) = \frac{(-1)^k}{\omega^k}\varphi^{(k)}\left(\frac{r}{\omega}z\right) = \frac{1}{2\pi}\int (is)^k e^{-s^2\omega^2/2}e^{-isr}\,ds, \tag{8}$$

where $\varphi(z) = (1/\sqrt{2\pi})\exp\{-z^2/2\}$ is the standard normal density and $H_k(z)$ are the Hermite polynomials, the first four of which are

$$H_0(z) = 1, \qquad H_1(z) = z, \qquad H_2(z) = z^2-1, \qquad H_3(z) = z^3-3z.$$

By inversion in this manner we find from (7) and (8) that the pdf of θ_n has the following asymptotic expansion as $n\to\infty$:

$$\mathrm{pdf}(r) = \frac{1}{\omega}\varphi\left(\frac{r}{\omega}\right)\left[1+\frac{1}{\sqrt{n}}\left\{a_1\left(\frac{r}{\omega}\right)+a_3\left(\frac{r}{\omega}\right)^3\right\}\right]+O(n^{-1}), \tag{9}$$

where

$$a_1 = -\frac{1}{2\omega}\,\mathrm{cgf}_{rs}\,e_{rs} - \frac{1}{2\omega^3}\,[3e_{rs}(\mathrm{cgf}_{rj}\,e_j)(\mathrm{cgf}_{sk}\,e_k) + in^{1/2}\mathrm{cgf}_{ijk}\,e_i\,e_j\,e_k], \quad \textbf{(10)}$$

$$a_3 = \frac{1}{2\omega^3}\,[3e_{rs}(\mathrm{cgf}_{rj}\,e_j)(\mathrm{cgf}_{sk}\,e_k) + in^{1/2}\mathrm{cgf}_{ijk}\,e_i\,e_j\,e_k], \quad \textbf{(11)}$$

$$\omega^2 = -\mathrm{cgf}_{ij}\,e_i\,e_j.$$

Expression (9) gives the Edgeworth expansion of the density of θ_n explicitly to order $O(n^{-1/2})$. Terms of higher order are obtained in an analogous way by carrying each of the expansions to higher level. They are given explicitly to order $O(n^{-1})$ in Sargan (1976, 1977a), although both sets of formulas have algebraic slips, and Phillips (1977). Withers (1983) subsequently gave correspondingly general formulas for the case of a regular functional of the empirical distribution of a random sample from an absolutely continuous distribution. The algebra quickly becomes impractical beyond the second order, but it is eminently suited to symbolic computing and numerical evaluation. The coefficients a_1 and a_3 determine the amount of bias and skewness in the distributional approximation, and their magnitude is determined by the second and third cumulants of the distribution of m^n (through the terms cgf_{rs} and cgf_{ijk}) and derivatives of θ_n to the second order (i.e., e_i and e_{rs}).

Three different levels of approximation are involved in the construction leading to (9):

(i) the approximation (1) of the statistic θ_n by a low-order polynomial in a vector of more basic moment statistics m^n;
(ii) the approximate representation (2) of the cf of θ_n in terms of the cf of m^n and its successive derivatives;
(iii) the approximate representation (5) and (6) of the cf of m^n in terms of the cf of the limiting distribution of m^n.

Of these approximations, (i) is often the most heroic. The adequacy of a polynomial representation of θ_n in terms of m^n will depend on many factors. The most important of these is the width of the domain in θ_n-space over which a good approximation to the distribution of θ_n is required. In many cases (e.g., when the dimension of m is large relative to n or when there are singularities in $\theta_n(m)$ close to the origin), it may not be possible to adequately capture the random behavior of θ_n by a polynomial approximation in anything else but a very small neighborhood of the origin (the probability limit of m). In other cases (e.g., when the moments m arise from time series whose parameters are close to the region of nonstationarity) the approximation (iii) may also be quite poor. When any of the approximations (i)–(iii) is inadequate, we can expect the Edgeworth expansion to perform poorly. Often, the worst affected regions of the distribution are the tails, where the shape of the true density necessarily depends on the behavior of θ_n in outlying regions of its space of definition.

In this connection, it is worth mentioning that whereas θ_n itself may possess no finite integral order moments and its distribution may have heavy tails, all moments of the approximating density in (9) are finite. In such cases, the finite moments of the approximating polynomial representation of θ_n become pseudomoments of θ_n. They can readily be calculated from (9). For instance, the first pseudomoment is given by

$$E(\theta_n) = \int \frac{1}{\omega}\varphi\left(\frac{r}{\omega}\right)\left[r + \frac{\omega}{\sqrt{n}}\left\{a_1\left(\frac{r}{\omega}\right)^2 + a_3\left(\frac{r}{\omega}\right)^4\right\}\right]dr = \frac{\omega}{\sqrt{n}}(a_1 + 3a_3)$$
$$= -\frac{1}{2\sqrt{n}}e_{rs}\text{cgf}_{rs} = -\frac{1}{2\sqrt{n}}e_{rs}\text{cf}_{rs}(0; m^n). \quad \textbf{(12)}$$

These moment approximations are sometimes called Nagar approximations, following their early use by Nagar (1959) for approximating the distribution of k-class estimates in simultaneous equations. Sargan (1974) developed criteria for the validity of such approximations, and Appendix A of his Walras–Bowley lecture (Sargan, 1976) extended that theory to a general setting comparable to the preceding situation, allowing for arbitrary order moments and expansions of arbitrary order. Appendix C (part 4) of the same lecture (Sargan, 1976) went on to apply the formulas (essentially (12) given previously) to extract the bias of the three-stage least squares estimator of a structural coefficient matrix in a simultaneous equations system. The same appendix utilized commutation matrix algebra to simplify the matrix manipulations in those derivations.

In later work, Sargan (1982) used pseudomoment expansions of the type given in (12) to help interpret the descriptive moment statistics conventionally reported in Monte Carlo experiments. When the moments of the underlying distribution are infinite, Sargan showed that such simulation-based moment statistics can be validly interpreted as estimates of the actual moments of the Edgeworth approximating distributions up to a certain order, depending on the sample size and the number of replications.

Many econometric test criteria are asymptotically chi-square rather than asymptotically normal. Expansions of the distributions of such criteria about their chi-square limits are also possible and of practical interest. Sargan (1980b) developed a validity theory for such expansions in a general functional context like that discussed previously but in the somewhat simpler case where some of the constituent moment functions m^n were normally distributed. An algorithm for extracting the explicit form of such expansions comparable in generality to the preceding material was given in Phillips and Park (1988). These results have proved useful in analyzing the behavior of econometric test statistics and comparing tests. But, like the expansion (9), they have not been found sufficiently reliable to implement in empirical research.

All of these contributions show Sargan's concern to produce methods that would be useful in practical econometric work and his desire to achieve general results. He clearly had a passion for finite sample theory and was equally comfortable doing exact theory, finding asymptotic expansions, proving validity theorems, or deriving algorithms. As it has transpired, none of these methods explored by Sargan are now in systematic use in practical work. They are also less frequently referenced in theoretical contributions than they were in the 1970s or early 1980s when the trail of finite sample theory in econometrics was being blazed. Nonetheless, they provide a substantial body of knowledge and technical innovation that is an important part of the fabric of this discipline.

5. CONCLUSION

The history of econometrics in the second half of the twentieth century is still to be written, but there is no doubt that Denis Sargan will figure in that history as one of the preeminent econometric theorists of the era. Sargan initiated a research agenda of great scope that influenced almost every major area of the discipline. His early work enriched and deepened the theory of instrumental variables and revealed the vast potential of this approach. His work on finite sample theory gave that burgeoning field both vitality and generality, offering the promise of empirical implementation and signaling the important role that computer simulation methods would play in improving inferential accuracy in applied research. His work on dynamic specification opened up a new field of research in the discipline that continues to impact applied time series research and thinking about econometric methodology. His research on dynamic panel modeling and unit roots pointed the way to new possibilities as those fields were beginning to emerge. At his retirement dinner held at Oxford University in 1985, Sargan told an enthralled audience of econometricians that he started off his academic career in econometrics with a few ideas that he wanted to pursue in econometric research and hoped to publish in good journals. That work was now pretty much completed, he said, and he was happy to hand over to a younger generation. One can hardly imagine a more modest way of summing up such a distinguished career.

Since Denis Sargan's passing, the world of econometric theory and its applications has moved on. But many of the themes of his research are present in ongoing work, albeit with new concerns occupying theorists and new approaches being followed by empirical researchers. Fortunately, Sargan's thinking about econometrics is preserved in his scientific contributions and is being kept alive through his teachings and the strong collegial influence he had in the econometrics community of the United Kingdom. The academic world has a great need for people like Denis Sargan who are prepared to invest in the people around them and help others in their own academic struggles. His life and career are a powerful example to us all.

REFERENCES

Aitken, A.C. (1935) On least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh* 55, 42–48.

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. In B.N. Petrov & F. Csaki (eds.), *Second International Symposium on Information Theory*, pp. 267–281. Budapest: Akademiai Kiado.

Andrews, D.F. & J.E. Stafford (2000) *Symbolic Computation for Statistical Inference*. Oxford: Oxford University Press.

Andrews, D.W.K. (1991), Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Arellano, M., L.P. Hansen, & E. Sentana (1999) Underidentification? Manuscript, CEMFI, Madrid.

Chambers, J.M. (1967) On methods of asymptotic approximation for multivariate distributions. *Biometrika*, 54, 367–383.

Chao, J.C. & N.R. Swanson (2002) Consistent estimation with a large number of weak instruments. Working paper, University of Maryland.

Forchini G. & G.H. Hillier (2002) Conditional Inference for Possibly Unidentified Structural Equations. Mimeo, University of York.

Hall, P. (1992) *The Bootstrap and Edgeworth Expansion*. New York: Springer-Verlag.

Han, C. & P.C.B. Phillips (2002) GMM with Many Moment Conditions. Mimeo, University of Victoria.

Hendry, D.F. (2001) Achievements and challenges in econometric methodology. *Journal of Econometrics* 100, 7–10.

Hendry, D.F. & H.-M. Krolzig (1999) Improving on 'Data mining reconsidered' by K.D. Hoover and S.J. Perez. *Econometrics Journal* 2, 41–58.

Hendry, D.F. & H.-M. Krolzig (2001) *Automatic Econometric Model Selection*. London: Timberlake Consultants Press.

Hendry, D.F & H.-M. Krolzig (2002) New developments in automatic general-to-specific modelling. In B.P. Stigum (ed.), *Econometrics and the Philosophy of Economics*. Cambridge: MIT Press.

Jeffreys, H. (1961) *Theory of Probability*, 3rd ed. Oxford: Oxford University Press.

Kleibergen, F. (2000) Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression. Discussion paper TI2000-055/4, Tinbergen Institute.

Koopmans, T.C. & W. Hood (1953) The estimation of simultaneous linear economic relationships. Chapter 6, pp. 112–199, in T.C. Koopmans and W. Hood (eds.), *Studies in Econometric Method*, Cowles Commission Monograph 14. New Haven: Yale University Press.

Leamer, E.E. (1978) *Specification Searches: Ad Hoc Inferences with Nonexperimental Data*. New York: Wiley.

Moreira, M.J. (2001) A Conditional Likelihood Ratio Test for Structural Models. Mimeo, University of California Berkeley.

Nagar, A.L. (1959) The bias and moment matrix of the general k-class estimators of the parameters in simultaneous equations. *Econometrica* 27, 575–595.

Phillips, A.W.H. (1956) Some notes on the estimation of time-forms of reactions in interdependent dynamic systems. *Economica* 23, 99–113.

Phillips, A.W.H. (1958) Cybernetics and the regulation of economic systems. *Cahiers de l'Institut de Science Economique Appliquée*, Serie N, 2, 41–50.

Phillips, P.C.B. (1977) Approximations to some finite sample distributions associated with a first order stochastic difference equation. *Econometrica* 45, 463–485.

Phillips, P.C.B. (1982) Small Sample Distribution Theory in Econometric Models of Simultaneous Equations. Cowles Foundation Discussion paper 617.

Phillips, P.C.B. (1989) Partially identified econometric models. *Econometric Theory* 5, 181–240.

Phillips, P.C.B. (1992) Bayes Methods for Trending Multiple Time Series with an Empirical Application to the US Economy. Cowles Foundation Discussion paper 1025.

Phillips, P.C.B. (1994) Some exact distribution theory for maximum likelihood estimators of cointegration coefficients in error correction models. *Econometrica* 62, 73–94.

Phillips, P.C.B. (1995a) Automated forecasts of Asia-Pacific economic activity. *Asia-Pacific Economic Review* 1, 92–102.

Phillips, P.C.B. (1995b) Bayesian model selection and prediction with empirical applications. *Journal of Econometrics* 69: 289–331.

Phillips, P.C.B. (1996) Econometric model determination. *Econometrica* 64, 763–812.

Phillips, P.C.B. (1998) Forecasts of Asia-Pacific economic activity to 2001:4. *Asia Pacific Economic Review* 4, 49–86.

Phillips, P.C.B. (2003) Laws and limits of econometrics. *Economic Journal*, forthcoming.

Phillips, P.C.B. & J.Y. Park (1988) On the formulation of Wald tests of non-linear restrictions. *Econometrica* 56, 1065–1084.

Rissanen, J. (1978) Modeling by shortest data description. *Automatica* 14, 465–471.

Robinson, P.M. (1991) Automatic frequency domain inference on semiparametric and nonparametric models. *Econometrica* 59, 1329–1363.

Sargan, J.D. (1958) The estimation of economic relationships using instrumental variables. *Econometrica* 26, 393–415.

Sargan, J.D. (1959) The estimation of relationships with autocorrelated residuals by the use of the instrumental variables. *Journal of the Royal Statistical Society, Series B* 21, 91–105.

Sargan, J.D. (1964) Wages and prices in the United Kingdom: A study in econometric methodology. In. P.E. Hart, G. Mills, & J.K. Whitaker (eds.), *Econometric Analysis for National Economic Planning*, vol. 16 of Colston Papers, pp. 25–63. London: Butterworth.

Sargan, J.D. (1974) The validity of Nagar's expansion for the moments of econometric estimators. *Econometrica* 42, 169–176.

Sargan, J.D. (1976) Econometric estimators and the Edgeworth expansion. *Econometrica* 44, 421–448.

Sargan, J.D. (1977a) Erratum. *Econometrica* 45, 272.

Sargan, J.D. (1977b) Some Tests for Dynamic Specification. Discussion paper A16, London School of Economics.

Sargan, J.D. (1980a) Some tests of dynamic specification for a single equation. *Econometrica* 48, 879–897.

Sargan, J.D. (1980b) Some approximations to the distribution of econometric criteria which are asymptotically distributed as chi-squared. *Econometrica* 48, 1107–1138.

Sargan, J.D. (1982) On Monte Carlo estimates of moments that are infinite. In R.I. Basmann & G.F. Rhodes (eds.), *Advances in Econometrics: A Research Annual*, pp. 267–299. Greenwich, CT: JAI Press.

Sargan, J.D. (1983a) Identification and lack of identification. *Econometrica* 51, 1605–1633.

Sargan, J.D. (1983b) Identification in models with autoregressive errors. In S. Karlin, T. Amemiya, & L.A. Goodman (eds.), *Studies in Econometrics, Time Series, and Multivariate Statistics*, pp. 169–205. New York: Academic Press.

Sargan, J.D. (1988a) *Contributions to Econometrics*, E. Maasoumi (ed.), vol. 1. Cambridge: Cambridge University Press.

Sargan, J.D. (1988b) *Contributions to Econometrics*, E. Maasoumi (ed.), vol. 2. Cambridge: Cambridge University Press.

Sargan, J.D. (2001) The choice between sets of regressors. *Econometric Reviews* 20, 159–170.

Schwarz, G. (1978) Estimating the dimension of a model. *Annals of Statistics* 6, 461–464.

Staiger, D. & J.H. Stock (1997) Instrumental variables regression with weak instruments. *Econometrica* 65, 557–586.

Wallace, D.L. (1958) Asymptotic approximations to distributions. *Annals of Mathematical Statistics* 29, 635–654.

Withers, C.S. (1983) Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals. *Annals of Statistics* 11, 577–587.