(Important contributors to this school of thought include Butterworth, 1972; Feltham, 1972; Ijiri, 1975; Beaver, 1998; and Christensen and Demski, 2002.)

The advantage of this view is it forces us to think in terms of complements and substitutes when dealing with this vast array of sources, and to look for economic forces that drive the disparity that bedevils the measurement school. And it is here that the comparative advantage of the accounting channel comes into focus: it is purposely designed and managed so that it is difficult to manipulate (Ijiri, 1975). This is why it often resorts to historical-cost measurement, as this removes major elements of subjectivity and manipulation potential. It is also why, in organized financial markets, most valuation information arrives before the firm's financial reports; and in this sense the financial reports provide a veracity check on the earlier reporting sources. In addition, cost allocation now enters as a natural phenomenon, either as a simple scaling device or – to use an analogy with informationally efficient markets – as a cousin to an information-based pricing kernel in a financial market (Christensen and Demski, 2002; Ross, 2004).

Libraries are organized in coordinated fashion, as are phone books; and the same can be said about accounting. A curiosity is the political side of the regulatory apparatus. It is difficult, for example, for the incumbent government to alter a government-provided statistical series, yet it is routine for the incumbent government to intervene in the accounting regulatory process. A second curiosity is the seemingly episodic nature of financial reporting frauds (Demski, 2003), although at the micro level it is well understood that opportunistic reporting is part of the game. For example, an ability to shift income from a later to an earlier period may be an inexpensive signal or, to speak more cynically, less costly to the firm than shifting real resources.

The disadvantage of the information school is its sheer breadth. The institutional context includes a vast array of information sources and actors, and sorting out first-order effects remains problematic.

Conclusion

Accounting, then, is simultaneously an important source of economic data and a collection of institutional regularities that provide research economists with yet another venue for documentation and exploration of economic forces. Why do we see episodic regulatory interventions? Why do we see forecasts of forthcoming accounting measures? Why do we not see supplementary estimation of economic depreciation? Why do we see the mix of historical-cost and market values that characterize modern financial reporting? Questions of this sort motivate much of the current research in accounting and finance.

JOEL S. DEMSKI

See also **assets and liabilities; capital measurement; cost functions; depreciation; double-entry bookkeeping; human capital; measurement; pensions; present value.**

Helpful comments by Haijin Lin and David Sappington are gratefully acknowledged.

Bibliography

Anderson, S., Hesford, J. and Young, M. 2002. Factors influencing the performance of activity based costing teams: a field study of ABC model development time in the automobile industry. *Accounting Organizations and Society* 27, 195–211.

Ball, R. and Brown, P. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6, 159–78.

Beaver, W. 1998. *Financial Reporting: An Accounting Revolution*. Englewood Cliffs, NJ: Prentice-Hall.

Butterworth, J. 1972. The accounting system as an information function. *Journal of Accounting Research* 10, 1–27.

Canning, J. 1929. *The Economics of Accountancy*. New York: Ronald Press.

Chambers, R. 1966. *Accounting, Evaluation and Economic Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Christensen, J. and Demski, J. 2002. *Accounting Theory: An Information Content Perspective*. New York: McGraw-Hill/Irwin.

Clark, J. 1923. *Studies in the Economics of Overhead Costs*. Chicago: University of Chicago Press.

Debreu, G. 1959. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*. New Haven, CT: Yale University Press.

Demski, J. 2003. Corporate conflicts of interest. *Journal of Economic Perspectives* 17(2), 51–72.

Edwards, E. and Bell, P. 1961. *The Theory and Measurement of Business Income*. Berkeley: University of California Press.

Feltham, G. 1972. *Information Evaluation*. Sarasota, FL: American Accounting Association.

Hicks, J. 1946. *Value and Capital*. Oxford: Clarendon Press.

Ijiri, Y. 1975. *Theory of Accounting Measurement*. Sarasota, FL: American Accounting Association.

Krantz, D., Luce, R., Suppes, P. and Tversky, A. 1971. *Foundations of Measurement*. New York: Academic Press.

Mock, T. 1976. *Measurement and Accounting Information Criteria*. Sarasota, FL: American Accounting Association.

Morgenstern, O. 1965. *On the Accuracy of Economic Observations*. Princeton, NJ: Princeton University Press.

Paton, W. 1922. *Accounting Theory*. New York: Ronald Press.

Petrick, K. 2002. Corporate profits: profits before tax, profits tax liability, and dividends. Methodology Paper. Washington, DC: Bureau of Economic Analysis. Online. Available at http://www.bea.gov///bea/ARTICLES/NATIONAL/NIPA/Methpap/methpap2.pdf, accessed 18 November 2005.

Ross, S. 2004. *Neoclassical Finance*. Princeton, NJ: Princeton University Press.

Solomons, D. 1965. *Divisional Performance: Measurement and Control*. New York: Financial Executives Research Foundation, 1965.

Wilson, R. 1983. Auditing: perspectives from multiperson decision theory. *Accounting Review* 58, 305–18.

adaptive estimation

An adaptive estimator is an efficient estimator for a model that is only partially specified.

For example, consider estimating a parameter that describes a sample of observations drawn from a distribution F. One natural question is: is it possible that an estimator of the parameter constructed without knowledge of F could be as efficient (asymptotically) as any well-behaved estimator that relies on knowledge of F? For some problems the answer is 'yes', and the estimator that is efficient is termed an adaptive estimator.

Consider the familiar scalar linear regression model (in which we let t rather than i index observations)

$$Y_t = \beta_0 + \beta_1 X_t + U_t,$$

where the regressor is exogenous and $\{U_t\}$ is a sequence of n independent and identically distributed random variables with distribution F. The parameter vector $\beta = (\beta_0, \beta_1)'$ is often of interest rather than the distribution of the error, F. If we assume that F is described by a parameter vector λ (that is, we parameterize the distribution), then the resultant (maximum likelihood or ML) estimator of β is parametric. If we assume only that F belongs to a family of distributions, then the resultant estimator of β is *semi*parametric. Because the OLS estimator does not require that we parameterize F, the OLS estimator is semiparametric. If the population error distribution is Gaussian, we know that the OLS estimator is equivalent to the ML estimator, and so is efficient. Although the OLS estimator is generally inefficient if F is not Gaussian, it may be possible to construct an alternative (semiparametric) estimator that retains asymptotic efficiency if F is not Gaussian. If we find that, for a family of distributions that includes the Gaussian, this estimator is asymptotically equivalent to the ML estimator, then this estimator is adaptive for that family.

The question then is: how can we verify that an estimator is adaptive? As there will generally be an arbitrarily large number of distributions in the family, it is not feasible to algebraically verify asymptotic equivalence for each distribution. In a creative paper, Stein (1956) first proposed a solution to this problem. Let $\{F_\lambda, \lambda \in \Lambda\}$ define a subset of the family of distributions, each member of which is parameterized by a value of λ (each member of this family must satisfy certain technical conditions, such as absolute continuity, which will not be explicitly defined). Although primary interest centers on β, the full set of parameters includes λ. The information matrix, evaluated at the population parameter values, is

$$\mathscr{I} = \begin{pmatrix} \mathscr{I}_{\beta\beta} & \mathscr{I}_{\beta\lambda} \\ \mathscr{I}_{\lambda\beta} & \mathscr{I}_{\lambda\lambda} \end{pmatrix},$$

where $\mathscr{I}_{\beta\beta}$ corresponds to the elements of β. Estimators of β (again, the estimators must satisfy technical conditions, such as $\sqrt{n}$ consistency, which are also not explicitly defined) will have covariance matrix that is at least as large as $\mathscr{I}^{\beta\beta}$, which is the upper left component of $\mathscr{I}^{-1}$. If the partial derivative of the log-likelihood with respect to β (the score for β) is orthogonal to the score for λ, then $\mathscr{I}_{\beta\lambda} = 0$ and $\mathscr{I}^{\beta\beta} = \mathscr{I}^{-1}_{\beta\beta}$. Because $\mathscr{I}_{\beta\beta}$ corresponds only to the parameter β, the asymptotically efficient estimator of β can be constructed without knowledge of λ. Stein argued that, if the condition $I_{\beta\lambda} = 0$ holds for all the elements of $\{F_\lambda\}$, then β is adaptively estimable.

While Stein's condition has intuitive appeal, it is not straightforward how to use the condition to define estimators that are adaptive. In an invited lecture, Bickel (1982) laid out a simpler condition that does yield a straightforward link to the construction of adaptive estimators. To understand the condition, let E_F denote expectation with respect to the population error distribution and let $E_{\tilde{F}}$ denote expectation with respect to an arbitrary distribution $\tilde{F} \in \mathscr{F}$. Let l be the log-likelihood for the regression model with data $z = (y, x)$ and let $\dot{l}(z, \beta, F)$ denote the score for β, constructed from the model in which F is the error distribution. A familiar condition that arises in the context of likelihood estimation is that the expected population score $E_F[\dot{l}(z, \beta, F)]$ equal 0. Bickel's condition is simply that the population score must have expectation zero over the entire family $\mathscr{F}$, that is, for any $\tilde{F} \in \mathscr{F}$,

$$E_{\tilde{F}}\left[\dot{l}(z, \beta, F)\right] = 0.$$

The two conditions are linked: if $\mathscr{F}$ is a convex family, then Stein's condition is implied by Bickel's condition. In detail, if $\mathscr{F}$ is a convex family, then $F_\lambda = \lambda F + (1 - \lambda)\tilde{F}$ with λ an element of $\Lambda = (0, 1)$. Bickel's condition then arises from Stein's condition by taking the limit as $\lambda \to 0$. For the linear regression model, an adaptive estimator of β exists for the family $\mathscr{F}$ that consists of all distributions that are symmetric about the origin (and several other technical conditions). If interest centres on the slope coefficient alone, then one need not restrict attention to distributions that are symmetric about the origin, as an adaptive estimator of β_1 can exist even if β_0 is not identified.

Bickel's score condition leads naturally to estimators that contain nonparametric estimators of the distribution,

$\hat{F}$. In consequence, adaptive estimation requires a second condition: the nonparametric estimator of the score must converge in quadratic mean to the population score. The resulting estimators of β are two-step estimators. The estimators require, as the first step, a $\sqrt{n}$-consistent estimator such as the OLS estimator. To understand the estimator's form, note that, if the distribution were known, then the two-step (linearized likelihood) estimator is

$$\hat{\beta}_{OLS} + n^{-1} \sum_{t=1}^{n} s\left(Z_t, \hat{\beta}_{OLS}, F\right),$$

with $s(Z_t, \hat{\beta}_{OLS}, F) = \mathscr{I}^{11}(\hat{\beta}_{OLS}, F)\dot{l}(Z_t, \hat{\beta}_{OLS}, F)$. The linearized likelihood estimator is asymptotically efficient. To form an adaptive estimator of β, we must replace F with a nonparametric estimator $\hat{F}$. If $\hat{F}$ is constructed so that $s(Z_t, \hat{\beta}_{OLS}, \hat{F})$ converges in quadratic mean to $s(Z_t, \hat{\beta}_{OLS}, F)$, then

$$\hat{\beta}_{AD} = \hat{\beta}_{OLS} + n^{-1} \sum_{t=1}^{n} s\left(Z_t, \hat{\beta}_{OLS}, \hat{F}\right)$$

is an adaptive estimator of β for the family $\mathscr{F}$.

For the linear regression model, as for numerous other models, nonparametric estimation of F entails nonparametric estimation of the density f. One popular nonparametric density estimator is the kernel estimator, which is employed by Portnoy and Koenker (1989) in their proof that semiparametric quantile estimators are also adaptive for β. If $\{\hat{U}_t\}$ denotes the OLS residuals, then a kernel density estimator is defined for all u in a small neighbourhood of each value of $\hat{U}_t$ as

$$\hat{f}_t(u) = (n-1)^{-1} \sum_{\substack{s=1 \\ s \neq t}}^{n} \xi_\sigma\left(u - \hat{U}_s\right),$$

where ξ_σ is a weight function that depends on the smoothing parameter σ. In Steigerwald (1992), ξ_σ corresponds to a Gaussian density with mean 0 and variance σ^2. The variance controls the amount of smoothing; as σ^2 declines, the weight given to residuals that lie some distance from $\hat{U}_t$ tends to zero. Of course, there are many other ways to form the nonparametric score estimator. Newey (1988) approximates the score by a series of moment conditions, which arise from exogeneity of the regressor and symmetry of F. Faraway (1992) uses a series of spline functions to approximate the score. Chicken and Cai (2005) use wavelets to form the basis for nonparametric estimation of f.

Recent results in adaptive estimation have focused on problems in which the error distribution is known, but other features are modelled nonparametrically. Some of the most intriguing results concern the type of stochastic differential equation often encountered in financial models. The price of an asset that is measured continuously over time, P_t, is often modelled as

$$dP_t = m_t dt + v_t dB_t.$$

The presence of standard Brownian motion, B_t, makes the model of price a stochastic differential equation. The function m_t captures the deterministic movement or drift while v_t is the potentially time-varying scale of the random component. Lepski and Spokoiny (1997) study the model in which v_t is constant and m_t is unknown. They establish that a nonparametric estimator of m is pointwise adaptive. Yet an estimator that is pointwise adaptive – that is, for a given point t_0 the nonparametric estimator of $m(t_0)$ is asymptotically efficient – may not perform well for all values within the range of the function m. Such an idea is intuitive; without knowledge of the smoothness of m, estimators designed to be optimal for one value of t may be very different from optimal estimators for another value of t. Cai and Low (2005) study efficient estimation of m over neighbourhoods of t_0 and show that an estimator constructed from wavelets is adaptive. The restriction that the scale is constant is often difficult to support with financial data. A more realistic model, which Mercurio and Spokoiny (2004) study, models the asset return as a stochastic differential equation with drift 0 and v_t varying over time. The time-varying scale is assumed to be constant over (short) intervals of time, but is otherwise unspecified. They construct a nonparametric estimator of the volatility from a kernel that performs local averaging and show that the resultant estimator is adaptive.

DOUGLAS G. STEIGERWALD

See also **efficiency bounds; partial linear model; semiparametric estimation.**

Bibliography

Bickel, P. 1982. On adaptive estimation. *Annals of Statistics* 10, 647–71.

Cai, T. and Low, M. 2005. Nonparametric estimation over shrinking neighborhoods: superefficiency and adaptation. *Annals of Statistics* 33, 184–213.

Chicken, E. and Cai, T. 2005. Block thresholding for density estimation: local and global adaptivity. *Journal of Multivariate Analysis* 95, 76–106.

Faraway, J. 1992. Smoothing in adaptive estimation. *Annals of Statistics* 20, 414–27.

Lepski, O. and Spokoiny, V. 1997. Optimal pointwise adaptive methods in nonparametric estimation. *Annals of Statistics* 25, 2512–46.

Mercurio, D. and Spokoiny, V. 2004. Statistical inference for time-inhomogeneous volatility models. *Annals of Statistics* 32, 577–602.

Newey, W. 1988. Adaptive estimation of regression models via moment restrictions. *Journal of Econometrics* 38, 301–39.

Portnoy, S. and Koenker, R. 1989. Adaptive L-estimation for linear models. *Annals of Statistics* 17, 362–81.

Steigerwald, D. 1992. On the finite sample behavior of adaptive estimators. *Journal of Econometrics* 54, 371–400.

Stein, C. 1956. Efficient nonparametric testing and estimation. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, ed. J. Neyman. Berkeley: University of California Press.

adaptive expectations

The adaptive expectations hypothesis may be stated most succinctly in the form of the equation:

$$E_t x_{t+1} = \sum_{i=0}^{\infty} \lambda(1-\lambda)^i x_{t-i}; \qquad 0 < \lambda < 1 \tag{1}$$

where E denotes an expectation, x is the variable whose expectation is being calculated and t indexes time. What this says is that the expectation formed at the present time, E_t of some variable, x, at the next future date, $t+1$, may be viewed as a weighted average of all previous values of the variable, x_{t-i}, where the weights, $\lambda(1-\lambda)^i$, decline geometrically. The weight attaching to the most recent, or current, observation is λ The above equation can be manipulated readily to deliver:

$$E_t x_{t+1} = E_{t-1} x_t + \lambda(x_t - E_{t-1} x_t). \tag{2}$$

What this equation says is that, viewed from time t, the expected value of the variable, x at $t+1$, is equal to the value which, at time $t-1$ was expected for t, plus an adjustment for the extent to which the variable turned out to be different at t from the value which, viewed from date $t-1$, had been expected. The change in the expectation is simply the fraction λ multiplied by the most recently observed forecast error. In this formulation, the adaptive expectations hypothesis is sometimes called the error learning hypothesis (see Mincer, 1969, pp. 83–90).

The adaptive expectations hypothesis was first used, though not by name, in the work of Irving Fisher (1911). The hypothesis received its major impetus, however, as a result of Phillip Cagan's (1956) work on hyperinflations. The hypothesis was used extensively in the late 1950s and 1960s in a variety of applications. L.M. Koyck (1954) used the hypothesis, though not in name, to study investment behaviour. Milton Friedman (1957), used it as a way of generating permanent income in his study of the consumption function. Marc Nerlove (1958) used it in his analysis of the dynamics of supply in the agricultural sector. Work on inflation and macro-economics in the 1960s was dominated by the use of this hypothesis. The most comprehensive survey of that work is provided by David Laidler and Michael Parkin (1975).

The adaptive expectations (or error learning) hypothesis became popular and was barely challenged from the middle-1950s through the late-1960s. It was not entirely unchallenged but it remained the only extensively-used proposition concerning the formation of expectations of inflation and a large number of other variables for something close to two decades. In the 1970s the hypothesis fell into disfavour and the rational expectations hypothesis became dominant.

The adaptive expectations hypothesis became and remained popular for so long for three reasons. First, in its error learning form it had the appearance of being an application of classical statistical inference. It looked like classical updating of an expectation based on new information.

Second, the adaptive expectations hypothesis was empirically easy to employ. Koyck (1954) showed how a simple transformation of an equation with an unobservable expectation variable in it could be rendered observable by performing what became a famous transformation bearing Koyck's name. If some variable, y, is determined by the expected future value of x, that is:

$$y_t = \alpha + \beta E_t x_{t+1} \tag{3}$$

where α and β are constants, then we can obtain an estimate of α and β by using a regression model in which equation (1) [or equivalently (2)] is used to eliminate the unobservable expected future value of x. To do this, substitute (1) into (3). Then write down an equation identical to (3) but for one period earlier. Multiply that second equation by $1-\lambda$ and subtract the result from (3) (Koyck, 1954, p. 22), to give:

$$y_t = \alpha\lambda + \beta\lambda x_t + (1-\lambda) y_{t-1} \tag{4}$$

An equation like this may be used to estimate not only the desired values of α and β but also the value of λ, the coefficient of expectations adjustment. Thus, economists seemed to have a very powerful way of modelling situations in which unobservable expectational variables were important and of discovering speeds of response both of expectations to past events and of current events to expectations of future events.

Third, the adaptive expectations hypothesis seemed to work. That is, when equations like (4) were estimated in the wide variety of situations in which the hypothesis was applied (see above), 'sensible' parameter values for α, β, λ were obtained and, in general, a high degree of explanatory power resulted.

If the adaptive expectations hypothesis was so intuitively appealing, easy to employ, and successful, why was

UDX

Request Date: 28-JUL-2012
Expiration Date: 06-AUG-2012
ILL Number:
ILL Number: 5321343
Call Number: HB61 N49 2008
Format: Part of Book
Title: New Palgrave Dictionary of Economics, 2nd Edition
Article Title: Adaptive Estimation
Volume/Issue: 1
Part Pub. Date: 2008
Pages: 13-14
Borrower: UCSB Library

Hum/SS Ref.

Printed Date: 31-JUL-2012
TGQ or OCLC #:
TGQ or OCLC #: 5320526
ID: USB1
ISBN/ISSN:
Publisher: Palgrave Macmillan
Address: UCSB Interlibrary Loan/Davidson Library/525 UCEN Road/Santa Barbara, CA, 93106/U.S.A.
UCSB Ariel IP 128.111.96.251; Email ill@library.ucsb.edu

Patron Name: Tulley, Stephanie (Faculty)
Patron e-mail: stulley@library.ucsb.edu
Service Level: Normal - Full Search
Delivery Method: Electronic Mail
Request Notes: 2nd edition only
Need By:
Verification Source: MELVYL-UCLinks-sfx:citation
Supplier Reference:
Local request number:
Owned By: UCD Shields Library

Service Type: Copy non returnable
Max Cost: USD55
Payment Type: IFM
Copyright Info:
Requester Symbol:
Return To: INTERLIBRARY BORROWING
SHIELDS LIBRARY - UC DAVIS
100 NORTH WEST QUAD
DAVIS, CA
95616-5292