KUNGL.
VETENSKAPSAKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Information Department, P.O. Box 50005, SE-104 05 Stockholm, Sweden
Phone: +46 8 673 95 00, Fax: +46 8 15 56 70, E-mail: info@kva.se, Web site: www.kva.se

---

Advanced information on
the Bank of Sweden Prize in Economic Sciences
in Memory of Alfred Nobel, 2000

The Scientific Contributions of James Heckman and Daniel McFadden

Microeconometrics and microdata

Microeconometric research is concerned with empirical analysis of the economic behavior of individuals and households, such as decisions on labor supply, consumption, migration or occupational choice. Microeconometric methods are equally relevant in studies of individual firms, for example their production and employment decisions. Over the last several decades, significant breakthroughs in empirical microeconomic research have been triggered by innovations in microeconometric methods and by greater availability of new types of data. The raw material in microeconometric research is *microdata,* where the units of observation are individuals, households or firms. Microdata appear as cross-section data and, to an increasing degree, as longitudinal (panel) data.

While offering new means of testing economic hypotheses and estimating economic models, the analysis of microdata has also raised new econometric problems. This, in turn, has inspired methodological research in microeconometrics, which can be loosely defined as a collection of econometric methods for handling problems of model specification, estimating and testing that arise in the analysis of microdata. A hallmark of recent microeconometric research is the close interplay of applied work on substantive economic issues and theoretical work on methodological problems.

New data bases have been crucial in this development. Until the late 1960s, the availability of data sources for empirical studies of individual economic behavior was very limited. Nowadays, there are a number of longitudinal data sets covering individuals and households, in the United States as well as in most European countries. A pathfinding early effort to create an infrastructure for microeconometric research was the Panel Study of Income Dynamics (PSID) set up by James Morgan and others at the University of Michigan in the late 1960s. The PSID has been used intensively in applied research and has served as a model for the construction of longitudinal data sets in other countries.

It is only recently that this remarkable growth of microdata on individuals and households has been matched by similar information on individual firms. As a result, microeconometric applications have largely been dominated by studies of individual and household behavior.

Microeconometric applications cover a wide range of fields in economics. Labor economists have used microeconometric techniques to study labor supply decisions, individual earnings, educational choice, worker mobility, and the duration of spells of employment and unemployment. Microeconometric methods are essential for studies in empirical public finance, e.g., the effects of taxes and welfare policies on labor supply; in consumer research, e.g., the choice of different brands; and in urban and transportation economics, e.g., the choice of residence or mode of transportation. Applied research in microeconomics and industrial organization relies on microeconometrics in studies of firms' production and factor demand decisions. Similar methods are also used by researchers in other social sciences.

James Heckman and Daniel McFadden have made fundamental contributions to microeconometrics. Heckman's most influential work deals with problems that arise when data are generated by a non-random selection process, a common phenomenon in microeconometric studies. McFadden's foremost contributions concern theory and methods for discrete choice analysis, such as the choice of occupation or mode of transportation.

James J. Heckman

James Heckman was born in Chicago, IL in 1944. After undergraduate studies at Colorado College, majoring in mathematics, he went on to study economics at Princeton University where he received his Ph.D. in 1971. Heckman has taught at Columbia University, Yale University and the University of Chicago. Since 1995 he is Henry Schultz Distinguished Service Professor of Economics at the University of Chicago.

Heckman's numerous contributions to microeconometric theory have been developed in conjunction with applied empirical research, especially in labor economics. His applied research covers labor supply, labor earnings, unemployment duration, evaluation of labor-market programs, fertility, and discrimination.

Heckman's analysis of selection bias in microeconometric research has profoundly changed applied research, in economics as well as in other social sciences. Selection bias may arise when a sample under study does not randomly represent the underlying population. The problem facing the analyst is to obtain estimates of relevant population parameters even in the wake of a selective sample. Non-random sample selection may result from individual decisions by the agents under study (*self-selection*), but may also reflect administrative rules, or decisions on the part of sampling statisticians.

Selection bias and self-selection

Selection problems are pervasive in applied microeconometric research. Working hours and wages are observed only for those individuals who have chosen to work; earnings of migrants are observed only for those who have chosen to move; earnings of university graduates are observed only for those who have completed a university education, and so on. The selection problem can be viewed as a problem of missing observations. Wages and hours cannot be observed among non-working individuals, had they chosen to work; likewise, the earnings of non-migrants, had they chosen to migrate, are unobservable to the analyst; analogously, there is a lack of information on the earnings of workers with a high-school education, had they pursued a university education.

Heckman's approach to the selection problem is closely linked to economic theory. His key insight is that observations are often missing because of conscious (self-selection) choices made by economic agents (e.g., the decision to work, to migrate or to pursue higher education). The relation between the reasons for missing observations and the nature of non-missing observations thus takes on an intriguing theoretical structure. Heckman's proposed solutions to selection problems can be appreciated not only statistically, but also in terms of microeconomic theory.

Heckman's contributions to the econometrics of selective samples emerged concurrently with his studies of labor supply in the mid-1970s. These studies pioneered the "second generation models" of labor supply, which are distinguished by estimating equations derived explicitly from utility maximization with stochastic error terms as an integral part of the model, rather than added as an afterthought. They enabled a unified analysis of the factors determining work hours and labor-force participation.

An important early example of this strand of research, the contribution in Heckman (1974) is remarkable for its treatment of the selectivity problem inherent in all studies of labor supply.[1] Standard economic theory views labor-force participation as a result of utility maximization, where the participants are individuals whose market wages exceed their reservation wages. To obtain unbiased estimates of basic structural parameters, the estimation procedure has to recognize the sample of labor-force participants is not the result of random selection, but the result of individual self-selection implied by utility maximization.

Heckman (1974) presented a model of married women's labor supply based on the utility maximization hypothesis. The sample of working women is self-selected in the sense that hours of work are only observed for women with market wages higher than their reservation wages. Heckman derived a likelihood function for this problem, estimated equations for market wages, the probability for working, and hours of work, and then used the estimated structural parameters to predict the probability of working, hours of work, reservation wages and market wages. This paper is an excellent example of how microeconomic theory can be combined with microeconometric methods to clarify an important economic issue.

Heckman's subsequent work offered computationally simpler methods for handling selection bias (Heckman 1976, 1979). The well-known Heckman correction – also called the two-stage method, Heckman's lambda or the Heckit method[2] – has become part of the standard toolbox in applied microeconometric work. The method may be described by means of the following two equations:

$$w_i = x_{1i}\beta_1 + \varepsilon_{1i} \ , \tag{1}$$

$$e_i^* = x_{2i}\beta_2 + \varepsilon_{2i} \ . \tag{2}$$

Equation (1) determines the individual's market wage, whereas (2) is a "participation equation" describing the individual's propensity to work. Thus, w_i is the observed market wage for individual i if she works and e_i^* a latent variable that

[1] See also Gronau (1974) and Lewis (1974) for important early discussions of self-selection in the context of data on wages and labor supply.

[2] The label "Heckit" was presumably invented to acknowledge similarities with the famous Tobit estimator due to 1981 economics laureate James Tobin (1958).

captures the propensity to work; x_{1i} and x_{2i} are vectors of observed explanatory variables, such as age and education; ε_{1i} and ε_{2i}, finally, are mean-zero stochastic errors representing the influence of unobserved variables affecting w_i and e_i^*. The parameters (vectors) of interest are β_1 and β_2.

Although the latent variable e_i^* is unobserved, we can define a dummy variable $e_i = 1$ if $e_i^* \geq 0$ and $e_i = 0$ otherwise; we thus observe the market wage only if $e_i = 1$, i.e., if the individual works. It is likely that the unobserved terms ε_{1i} and ε_{2i} are positively correlated; individuals with higher wages, given x_{1i} and x_{2i}, are presumably also more likely to work. If so, the sample of individuals observed as working will not accurately represent the underlying population, even in a large sample. Failure to recognize this selectivity generally produces inconsistent estimates of the parameters in the wage equation.

Heckman suggested a simple method to deal with this selection problem. Note that the conditional mean of ε_{1i} can be written as:

$$E(\varepsilon_{1i} \mid e_i^* \geq 0) = E(\varepsilon_{1i} \mid \varepsilon_{2i} \geq -x_{2i}\beta_2) \ , \tag{3}$$

and hence

$$E(w_i \mid x_{1i}, e_i = 1) = x_{1i}\beta_1 + E(\varepsilon_{1i} \mid \varepsilon_{2i} \geq -x_{2i}\beta_2) \ . \tag{4}$$

Thus, the regression equation on the selected sample depends on both x_{1i} and x_{2i} . Omitting the conditional mean of ε_{1i} biases the estimates of β_1 (unless ε_{1i} and ε_{2i} are uncorrelated, in which case the conditional mean of ε_{1i} is zero). Selection bias can thus be regarded as a standard problem of omitted-variable bias. The problem is to find an empirical representation of the conditional mean of ε_{1i} and include this variable in the wage equation.

Under the assumption that ε_{1i} and ε_{2i} are drawn from a bivariate normal distribution, we can derive the regression equation:

$$E(w_i \mid x_{1i}, e_i = 1) = x_{1i}\beta_1 + \rho\sigma_1\lambda_i \ . \tag{5}$$

In (5) ρ is the correlation coefficient between ε_{1i} and ε_{2i}, σ_1 is the standard deviation of ε_{1i}, and λ_i – the inverse of Mill's ratio – is given by

$$\lambda_i = \frac{\phi(x_{2i}\beta_2/\sigma_2)}{\Phi(x_{2i}\beta_2/\sigma_2)} \ , \tag{6}$$

where ϕ and Φ are the density and distribution functions of the standard normal distribution and σ_2 is the standard deviation of ε_{2i}.

Heckman showed how to estimate (5) in a two-step procedure. The first step involves estimating the parameters in (2) by the probit method, using the entire sample. These estimates can then be used to compute λ_i for each individual in the sample. Once λ_i is computed, we can estimate (5) over the sample of working individuals by ordinary least squares regression, treating $\rho\sigma_1$ as the regression coefficient for λ_i.

The sign of the selection bias depends on the correlation between the errors in the wage and participation equations (ρ) and the correlation between λ_i and the variables in the wage equation (x_{1i}). Since λ_i is a decreasing function of the probability of sample selection, it follows that the β-coefficient on variables in x_{1i} that are likely to raise both wages and participation, such as education, will be biased downwards if the Heckit technique is not applied (provided $\rho > 0$).

Heckman's seminal work in this area has generated many empirical applications in economics as well as in other social sciences and applied statistics dealing with non-randomly missing data. One early example of empirical applications is the paper by Lee (1978), who examined relative wage effects of union membership in the United States, recognizing that membership is not random but the outcome of individual self-selection. Another is Willis and Rosen (1979), who investigated the wage premiums associated with higher education, recognizing endogenous educational choice as depending on the perceived gains to education. Later on, Heckman and Guilherme Sedlacek (1985) presented an empirical model of the sectoral allocation of workers in the U.S. labor market along the lines of the Roy (1951) model of income distribution: utility-maximizing individuals can work in several sectors, but only in one sector at a time. The analysis includes an assessment of how self-selection impacts on wage inequality.

Heckman's work has also generated a sizable literature on econometric method. The original model has been extended in a number of ways by Hackman and others.[3] These efforts, typically aimed at eliminating the restrictive assumption of bivariate normality, have involved the use of semi-parametric methods; see, for example, Heckman and Robb (1985b), Heckman (1990), Manski (1989), Newey, Powell and Walker (1990) and Lee (1994a,b).

Evaluation of active labor-market programs

Active labor-market programs, such as training, job-search assistance and employment subsidies, have become increasingly widespread in most OECD countries. Such programs are generally targeted at individuals who are unemployed or have low skills or earnings.

The classical problem of a program evaluation is to determine how participation in a specific program affects individual outcomes, such as earnings or employment, compared to non-participation. The paramount difficulty is to characterize the counterfactual situation, i.e., to answer the question: what would have happened if the individual had not participated in the program? Since it is impossible to observe an individual as both participant and non-participant, it is necessary to use information on non-participants' outcomes for this purpose. Given that the allocation of individuals to programs is seldom purely random, the group of participants becomes a selected sample with observed and unobserved characteristics that may differ from those of the overall population.

James Heckman is the world's foremost researcher on econometric policy evaluation. In this area, as in his labor-supply research, Heckman relies on a structural approach based on microeconomic theory to guide the model specification and to interpret the empirical results. The main ingredients of policy

[3] See Vella (1998) for a survey of econometric literature following Heckman's seminal work.

evaluation are twofold: (i) a model of participation in programs, and (ii) a model of program outcomes. Heckman's research on program evaluation can be seen as a natural continuation of his earlier work on selection models. In joint work with others, Heckman has presented a number of new results concerning identification and estimation of the effects of social programs (Heckman and Robb, 1985 a,b) and, more recently, improved our understanding of the *pros* and *cons* of using experimental rather than non-experimental data in program evaluation (Heckman and Smith, 1995; Heckman, Ichimura and Todd, 1997; Heckman, Smith and Clements, 1997). The latter set of studies has also spurred theoretical advances in the use of matching methods as econometric evaluation estimators. On balance, Heckman does not emerge as a strong supporter of the experimental approach, arguing that social experiments are valid only under special statistical and behavioral assumptions. A main lesson seems to be that there are no universally correct methods for evaluating programs. What method works best depends on the issue in question and on the economic models that determine participation and outcomes. Heckman has also presented substantive empirical results on the effects of various labor-market programs. Conclusions regarding the effects are often somewhat pessimistic. A survey primarily of U.S. studies (Heckman, LaLonde and Smith, 1999) concludes that programs frequently have very small (sometimes negative) effects for the participants and do not appear to pass conventional cost-benefit tests. On the other hand, there does seem to be substantial heterogeneity in program outcomes across participants and types of programs.

Duration models

The analysis of duration data has a long tradition in engineering and biomedical research.[4] More recently, it has also entered into economic and social science research, where duration models have been applied to a variety of problems. Duration models are now standard tools when studying the length of unemployment spells, demographic events (marriage, fertility, mortality, and migration), political events (e.g., how the occurrence of government crises depends on the time elapsed since the last election), and some industrial relations (e.g., the length of strikes). The models are also used in consumer research, to study the timing of purchases of products, as well as in macroeconomic research, to examine issues such as the duration of business cycles.

In his work on econometric duration analysis, Heckman has been particularly preoccupied with the effects of unobserved heterogeneity, i.e., individual differences in unobserved variables that may influence the duration of unemployment or employment. As unobserved heterogeneity in the context of duration data introduces specific selection problems, Heckman's work in this area fits well with his overall research agenda on sample selection.

This may be exemplified by studies of how the exit rate from unemployment to employment evolves over a spell of unemployment. One problem here is that

[4] So-called failure time analysis (e.g., analysis of the durability of electrical equipment) has been common in engineering for decades. Survival analysis likewise has a long tradition in biomedical research (e.g., studies of survival after surgery).

individuals with relatively weak employment prospects seem to "survive" as unemployed to a higher degree than individuals with more favorable characteristics. Thus, the "quality" of the stock of unemployed at each point in time is the result of a selection process partly driven by factors unobserved by the analyst. Overrepresentation of more "unemployment-prone" individuals at long durations can easily lead to the conclusion of negative duration dependence, i.e., that the exit rate to employment declines over a spell of unemployment. However, what appears to be negative duration dependence may simply be a sorting or selection effect.

In joint work with Burton Singer, Heckman addressed the problem of treating unobserved heterogeneity without imposing restrictive assumptions regarding the distribution of unobserved variables. Heckman and Singer (1984a) proposed a non-parametric estimator that has become widely used in applied work in economics and demography.

Other noteworthy contributions by Heckman to the duration literature include identification results for a class of duration models (Heckman and Singer, 1984b and Heckman and Honoré, 1989) and further treatment of identification issues (Heckman, 1991 and Heckman and Taber, 1994). Heckman has also written applied empirical papers on unemployment duration and fertility.

Daniel L. McFadden

Daniel McFadden was born in Raleigh, NC in 1937. He received his undergraduate degree from the University of Minnesota, with a major in physics. McFadden switched to economics in the late 1950s and received a Ph.D. from the University of Minnesota in 1962. His academic appointments include professorships at the University of Pittsburg, Yale University, Massachusetts Institute of Technology and the University of California at Berkeley. Since 1990 he is E. Morris Cox Professor of Economics at Berkeley.

McFadden's most important and influential contribution is his development of the economic theory and econometric methodology for discrete choice, i.e., choice among a finite set of alternatives. He has also made significant contributions in other fields of economics, including production theory and environmental economics. McFadden's research may best be characterized by his ability to combine the development of theory and methodology with applications to substantive empirical problems.

Discrete choice analysis

Discrete choice problems appear frequently in economics as well as in other social sciences. Consider, for example, the modeling of phenomena such as individual labor-force participation, occupational or locational decisions, or travel mode choice. Here the observations to be explained are discrete (or qualitative) and cannot be represented by continuous variables. Standard demand theory as well as traditional econometric methods, aimed at explaining variations in continuous variables, are generally inappropriate to analyze discrete choice behavior.

Problems of qualitative choice were initially dealt with in work in psychometrics, biometrics and to some extent also econometrics. Early contributions of particular importance are Thurstone (1927) and Luce (1959), who formulated models of discrete probabilistic choice. According to the psychological interpretation, individual choice behavior is intrinsically probabilistic.

By contrast, the economic and econometric approach developed by McFadden treats individual choice as deterministic. It focuses instead on the lack of information on the part of the analyst, such as imperfect information about the characteristics of alternatives and individuals under study. Whereas psychologists have generally been concerned with individual choices *per se*, economists have generally been more interested in aggregate outcomes, such as the fraction of a population that selects a certain alternative.[5]

The conditional logit model

McFadden's most fundamental contribution is the integration of economic theory and econometric methodology for discrete choice analysis. His seminal paper, entitled "Conditional Logit Analysis of Qualitative Choice Behavior" (McFadden, 1974a), and contemporaneous empirical case studies fundamentally changed researchers' thinking about the econometric analysis of individual behavior. Discrete choice analysis rapidly developed into one of the main fields of modern econometrics.

McFadden's approach may be sketched as follows. Suppose that each individual in a population faces a finite set of alternatives and chooses an alternative that maximizes his or her utility. The data available to the analyst are assumed to be generated by the repeated drawing of an individual at random from the population and recording a vector a of the individual's attributes (such as age, gender, income, etc.), the set $\mathcal{I}$ of alternatives available to the individual (e.g., travelling by car, bus, subway, etc.), and the individual's actual choice i from the set $\mathcal{I}$. Assume that the individual's utility from choosing i is of the additive form $u(i,a) = v(i,a) + e(i,a,\omega)$, where $v(i,a)$ is common to all individuals with observed attributes a, while $e(i,a,\omega)$ is particular to the drawn individual ω. Both utility terms are deterministic, the first reflecting "representative" tastes in the population, and the second reflecting idiosyncratic taste variations. Treating the unobserved utility terms $e(i,a,\omega)$ as realizations of random variables $\varepsilon(i,a)$, and letting $P(i \mid a,\mathcal{I})$ denote the *conditional choice probability* that the randomly drawn individual will choose alternative $i \in \mathcal{I}$, given his or her observed attributes a and the set of alternatives $\mathcal{I}$, we obtain:

$$P(i|a,\mathcal{I}) = \Pr[v(i,a) + \varepsilon(i,a) \geq v(j,a) + \varepsilon(j,a) \quad \forall j \in \mathcal{I}] \ . \tag{7}$$

This is called the *additive random utility model (ARUM)* of discrete choice. The right-hand side of (7) is the probability that an individual drawn at random from

[5] See Ben-Akiva and Lerman (1985) and Anderson, de Palma and Thisse (1992) for surveys and discussions of discrete choice models.

the population has a utility function that makes i the utility-maximizing choice, given the individual's attributes a and choice set $\mathcal{I}$.

If the random vector $\langle \varepsilon(i,a) \rangle_{i \in \mathcal{I}}$ has a joint cumulative distribution function F, and we write $\mathcal{I} = \{1, ..., I\}$, the right-hand side of equation (7) can be written as an integral in the i'th partial derivative, F_i, of F:

$$P(i|a, \mathcal{I}) = \int_{-\infty}^{+\infty} F_i[x + v(i,a) - v(1,a), ..., x + v(i,a) - v(I,a)]dx \ . \qquad (8)$$

In particular, if the random variables $\varepsilon(i,a)$ are independently distributed with cumulative distribution function $\exp[e^{-\sigma x_i}]$ for $\sigma > 0$ (the Gumbel, or type-I, extreme-value distribution), their joint distribution F becomes:

$$F(x_1, ..., x_I) = \exp\left[-\sum_{j \in \mathcal{I}} e^{-\sigma x_j}\right], \qquad (9)$$

and the choice probabilities in equation (7) reduce to the analytically convenient logit form:

$$P(i \mid a, \mathcal{I}) = \frac{\exp \sigma v(i,a)}{\sum_{j \in \mathcal{I}} \exp \sigma v(j,a)} \ . \qquad (10)$$

The parameter $\sigma > 0$ is inversely proportional to the standard deviation of the random utility terms $\varepsilon(i,a)$. In the limit as $\sigma \to \infty$, the choice probabilities $P(i \mid a, \mathcal{I})$ in (10) assign all probability mass to the alternatives with maximum "representative" utility $v(i,a)$ — and we obtain the traditional microeconomic model of fully deterministic utility maximization.

In order to make the resulting logit model tractable for predictive purposes, it is usually assumed that the "representative" utility terms $v(i,a)$ depend on known characteristics of the alternatives and the population in some analytically tractable way. For example, in the case of travel mode choice, these characteristics could be travel time, travel costs, etc. The associated parameter vectors can then be estimated by the maximum likelihood method.

McFadden called his innovation the *conditional logit model*. Although multinomial logit models had been around for some time (Theil, 1969; Quandt, 1970), McFadden's derivation of the model based on an economic theory of population choice behavior was entirely new. His contribution was immediately recognized as a paradigmatic breakthrough, and paved the way for statistical estimation and applications.

Subsequent development of discrete choice analysis

The attractiveness of the multinomial logit model lies in its combination of solid microeconomic foundations and computational simplicity. This simplicity follows from the assumption of statistical independence of the random utility terms, an assumption which implies *independence of irrelevant alternatives* (IIA). The ratio of the probabilities of choosing any two alternatives is independent of the properties of all other alternatives, as can be noted by using (10) and taking the ratio of any two alternatives. For example, an expanded choice set does not affect the odds ratio for two choices. Under the IIA assumption, parameters pertaining to two alternatives can be consistently estimated by using data only on those individuals who have chosen these two alternatives. Thus, the IIA property allows estimation of the multinomial logit model on choice-based samples, which are much more easily obtained than population samples, with due compensation for the arising sample bias; see Manski and Lerman (1977) and Manski and McFadden (1981).

However, as McFadden has pointed out, the IIA assumptions is restrictive in many applications. For example, it is unlikely that the odds ratio of any two choices would be invariant to the introduction of a new alternative that is a close substitute for an existing alternative.[6] This problem exemplifies the more general difficulty of estimating population parameters from self-selected samples. Hausman and McFadden (1984) devised a procedure for testing the validity of the IIA assumption based on the idea of comparing estimates from a self-selected subset with estimates from the full choice set; the two estimates should not change systematically if the IIA assumption is valid. Further specification tests were developed in McFadden (1987).

McFadden has also showed how to relax the IIA assumption through his development of *nested multinomial logit* and *generalized extreme value (GEV) models.* The nested logit model, introduced by Ben-Akiva (1973) and McFadden (1978), relaxes the IIA assumption by permitting certain statistical dependence between the choices. In this model individuals' decisions can be interpreted as having a hierarchical structure. Consider, for instance, the joint choice of destination and travel mode. One possible nested logit formulation of this decision problem would be to assume that for each destination the individual selects the preferred mode of transportation and, taking this into account, chooses his destination.

The GEV model, developed by McFadden (1978, 1981), is more general and analytically elegant. To derive it, note that if we generalize equation (9) to

$$F(x_1, ..., x_I) = \exp\left[-G(e^{-\sigma x_1}, ..., e^{-\sigma x_I})\right] \ , \tag{11}$$

for some linearly homogeneous function G, equation (7) becomes

$$P(i \mid a, \mathcal{I}) = \frac{G_i(e^{-\sigma v(1,a)}, ..., e^{-\sigma v(I,a)}) \exp \sigma v(i,a)}{G(e^{-\sigma v(1,a)}, ..., e^{-\sigma v(I,a)})} \ , \tag{12}$$

where G_i is the partial derivative of G with respect to its i'th argument. Using Euler's formula, the denominator can be written as a sum over G's partial derivatives, and we obtain:

$$P(i \mid a, \mathcal{I}) = \frac{\exp \sigma v(i,a) + \ln G_i(e^{-\sigma v(1,a)}, ..., e^{-\sigma v(I,a)})}{\sum_j \exp[\sigma v(j,a) + \ln G_j(e^{-\sigma v(1,a)}, ..., e^{-\sigma v(I,a)})]} \ . \tag{13}$$

[6] The 1983 economics laureate Gerard Debreu (1960) was the first to point this out, in the context of Luce's (1959) probabilistic choice model.

In other words, the choice probabilities still have a logit form. Now, however, the choice probability for an alternative i depends not only on its own observed attributes, via $v(i,a)$, but also on the observed attributes of other alternatives, such that IIA need no longer hold. The usual logit model is obtained as a special case when the function G is the sum of its arguments (then all partial derivatives are equal to unity, and the second term in the numerator and denominator of (13) vanishes), and the nested logit model is obtained as a special case when G is a CES function.

Another useful generalization of the multinomial logit model is the so-called *mixed logit model*. This generalization is obtained by aggregating choice behaviors across diverse subpopulations, where all individuals in each subpopulation have the same observed attributes, and each subpopulation's choice behavior is modelled as outlined above. As shown by McFadden and Train (1998), any well-behaved random utility model of discrete choice can be approximated to any degree of accuracy by such a mixed logit model. Applications of this model usually require Monte Carlo simulation methods.

Yet another approach is the *multinomial probit model* with correlated errors. However, as with the mixed logit model, computational difficulties arise when this model is applied to problems with more than a few alternatives, mainly because the calculation of choice probabilities involves evaluating multiple integrals. Lerman and Manski (1981) introduced the idea of computing the choice probabilities by means of Monte Carlo simulation methods where repeated random draws are taken from a multivariate normal distribution. McFadden (1989) further developed this idea by proposing an estimation approach known as the *method of simulated moments*. McFadden's article resolved the basic statistical properties of this method and a sizeable literature has since emerged in this area.

A natural step in the evolution of individual choice analysis is the development of models and methods that explain both discrete and continuous choices. The article by Dubin and McFadden (1984) is noteworthy as a fine example of how to integrate a general methodological contribution with an empirical study of practical usefulness (the household's discrete choice among electric appliances and its continuous choice of energy consumption).

As already emphasized, most of McFadden's work is indeed characterized by a close relation between economic theory, econometric methodology and applied empirical studies. Early empirical applications to urban travel demand are thus reported in McFadden (1974b) and Domencich and McFadden (1975). During the 1980s, and 1990s, McFadden was engaged in empirical work on residential energy demand (Cowing and McFadden, 1984), the demand for telephone services (McFadden, Train and Ben-Akiva, 1987) and the demand for housing among the elderly (McFadden, 1994a).

Other contributions

McFadden has made important contributions in several other fields. In the 1960s and early 1970s, he worked intensively on the theoretical and econometric analysis of production. Most of these contributions remained unpublished until they appeared in the two-volume collection of papers edited by Fuss and McFadden (1978). McFadden's work in this area became highly influential and established the principle of duality between cost or profit functions and production functions as a principal tool in the empirical analysis of production.

In other important work, in collaboration with Peter Diamond, McFadden exploited the duality between the expenditure function and utility maximizing demand functions to explore such problems as the deadweight burden of taxation and optimal commodity taxes (Diamond and McFadden, 1974). This paper established duality as an indispensable tool in modern public economics.

In the 1990s, McFadden has contributed to environmental economics, studying the willingness-to-pay for natural resources. McFadden (1994b) examined in detail the properties of the contingent valuation method for estimating the so-called existence value of natural resources and developed new econometric techniques. With Jerry Hausman and Gregory Leonard, he developed an empirical discrete choice model to assess the welfare losses caused by natural resource damage (Hausman, Leonard and McFadden, 1995). The model was applied to recreational demand in Alaska and, in particular, to estimate the welfare losses to Alaskans caused by the oil spill from the tanker Exxon Valdez in 1989. McFadden's work in this field is yet another example of his masterly skills in integrating economic theory, econometric methodology and substantive empirical applications.

References

Anderson S., A. de Palma and J-F. Thisse (1992), *Discrete Choice Theory of Product Differentiation*, MIT Press.

Ben-Akiva M. (1973), Structure of Travel Passenger Demand Models, Ph. D. thesis, MIT.

Ben-Akiva M. and S. Lerman (1985), *Discrete Choice Analysis: Theory and Applications to Predict Travel Demand*, MIT Press.

Cowing T. and D. McFadden (1984), *Microeconomic Modeling and Policy Analysis: Studies in Residential Energy Demand*, Academic Press.

Debreu G. (1960), Review of D. Luce, *Individual Choice Behavior: A Theoretical Analysis*, *American Economic Review* 50, 186–188.

Diamond P. and D. McFadden (1974), Some Uses of the Expenditure Function in Public Finance, *Journal of Public Economics* 3, 3–21.

Domencich T. and D. McFadden (1975), *Urban Travel Demand: A Behavioral Analysis*, North-Holland.

Dubin J. and D. McFadden (1984), An Econometric Analysis of Residential Electric Appliance Holdings and Consumption, *Econometrica* 52, 345–362.

Fuss M. and D. McFadden (eds.) (1978), *Production Economics: A Dual Approach to Theory and Applications*, vols I & II, North-Holland.

Gronau R. (1974), Wage Comparisons - a Selectivity Bias, *Journal of Political Economy* 82, 1119–1143.

Hausman J. and D. McFadden (1984), Specification Tests for the Multinomial Logit Model, *Econometrica* 52, 1219–1240.

Hausman J., G. Leonard and D. McFadden (1995), A Utility-Consistent, Combined Discrete Choice and Count Data model: Assessing Recreational Use Losses due to Natural Resource Damage, *Journal of Public Economics* 56, 1–30.

Heckman J. J. (1974), Shadow Wages, Market Wages and Labor Supply, *Econometrica* 42, 679–693.

Heckman J. J. (1976), The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *Annals of Economic and Social Measurement* 5, 475–492.

Heckman J. J. (1979), Sample Selection Bias as a Specification Error, *Econometrica* 47, 153–161.

Heckman J. J. (1990), Varieties of Selection Bias, *American Economic Review* 80, 313–318.

Heckman J. J. (1991), Identifying the Hand of the Past: Distinguishing State Dependence from Heterogeneity, *American Economic Review* 81, 75–79.

Heckman J. J. and B. Honoré (1989), The Identifiability of the Competing Risks Model, *Biometrica* 89, 325–330.

Heckman J. J., H. Ichimura and P. Todd (1997), Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, *Review of Economic Studies* 64, 605–654.

Heckman J. J., R. LaLonde and J. Smith (1999), The Economics and Econometrics of Active Labor Market Programs, *Handbook of Labor Economics*, vol 3A, North-Holland.

Heckman J. J. and R. Robb (1985a), Alternative Methods for Evaluating the Impact of Interventions: An Overview, *Journal of Econometrics* 30, 239–267.

Heckman J. J. and R. Robb (1985b), Alternative Methods for Evaluating the Impact of Interventions, in J. J. Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data*, Econometric Society Monographs Series, Cambridge University Press.

Heckman J. J. and G. Sedlacek (1985), Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market, *Journal of Political Economy* 93, 1077–1125.

Heckman J. J. and B. Singer (1984a), A Method of Minimizing the Impact of Distributional Assumptions for Duration Data, *Econometrica* 52, 271–320.

Heckman J. J. and B. Singer (1984b), The Identifiability of the Proportional Hazard Model, *Review of Economic Studies*, 231–241.

Heckman J. J. and J. Smith (1995), Assessing the Case for Social Experiments, *Journal of Economic Perspectives* 9, 85–110.

Heckman J. J., J. Smith and N. Clements (1997), Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts, *Review of Economic Studies* 64, 487–535.

Heckman, J. J. and C. Taber (1994), Econometric Mixture Models and More General Models for Unobservables in Duration Analysis, *Statistical Methods in Medical Research* 3, 279–302.

Lee L-F. (1978), Unionism and Wage Rates: Simultaneous Equations Models with Qualitative and Limited Dependent Variables, *International Economic Review* 19, 415–433.

Lee L-F. (1994a), Semiparametric Instrumental Variables Estimation of Simultaneous Equation Sample Selection Models, *Journal of Econometrics* 63, 341–388.

Lee L-F. (1994b), Semiparametric Two-Stage Estimation of Sample Selection Models Subject to Tobit-Type Selection Rules, *Journal of Econometrics* 61, 305–344.

Lerman S. and C. Manski (1981), On the Use of Simulated Frequencies to Approximate Choice Probabilities, in C. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press.

Lewis H.G. (1974), Comments on Selectivity Biases in Wage Comparisons, *Journal of Political Economy* 82, 1145–1155.

Luce R. D. (1959), *Individual Choice Behavior: A Theoretical Analysis*, Wiley.

Manski C. (1989), Anatomy of the Selection Problem, *Journal of Human Resources* 24, 343–360.

Manski C. and S. Lerman (1977), The Estimation of Choice Probabilities from Choice Based Samples, *Econometrica* 45, 1977–1988.

Manski C. and D. McFadden (1981), Alternative Estimators and Sample Designs for Discrete Choice Analysis, in C. Manski and D. McFadden (eds.), *Structural Analysis and Discrete Data with Econometric Applications*, MIT Press.

McFadden D. (1974a), Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (ed.), *Frontiers of Econometrics*, Academic Press.

McFadden D. (1974b), The Measurement of Urban Travel Demand, *Journal of Public Economics* 3, 303–328.

McFadden D. (1978), Modelling the Choice of Residential Location, in A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull (eds.), *Spatial Interaction Theory and Planning Models*, North-Holland.

McFadden D. (1981), Econometric Models for Probabilistic Choice, in C. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, Harvard University Press.

McFadden D. (1987), Regression Based Specification Tests for the Multinomial Logit Model, *Journal of Econometrics* 34, 63–82.

McFadden D. (1989), A Method for Simulated Moments for Estimation of Discrete Response Models without Numerical Integration, *Econometrica* 57, 995–1026.

McFadden D. (1994a), Demographics, the Housing Market, and the Welfare of the Elderly, in D. Wise (ed.), *Studies in the Economics of Aging*, University of Chicago Press.

McFadden D. (1994b), Contingent Valuation and Social Choice, *American Journal of Agricultural Economics* 74, 689–708.

McFadden D. and K. Train (1998), Mixed MNL Models for Discrete Response, *Journal of Applied Econometrics*, forthcoming.

McFadden D., K. Train and M. Ben-Akiva (1987), The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices, *Rand Journal of Economics* 18, 109–123.

Newey W., J. Powell and J. Walker (1990), Semiparametric Estimation of Selection Models, *American Economic Review* 80, 324–328.

Quandt R. (1970), *The Demand for Travel: Theory and Measurement,* Heath.

Roy A. (1951), Some Thoughts on the Distribution of Earnings, *Oxford Economic Papers* 3, 135–146.

Theil H. (1969), A Multinomial Extension of the Linear Logit Model, *International Economic Review* 10, 251–259.

Thurstone L. (1927), A Law of Comparative Judgement, *Psychological Review* 34, 273–286.

Tobin J. (1958), Estimation of Relationships for Limited Dependent Variables, *Econometrica* 26, 24–36.

Vella F. (1998), Estimating Models with Sample Selection Bias: A Survey, *Journal of Human Resources* 33, 127–169.

Willis R. and S. Rosen (1979), Education and Self-Selection, *Journal of Political Economy* 87, S1–S36.