Journal of Economic Literature
Vol. XXXVI (December 1998), pp. 2108–2165

Empirical Approaches to the Measurement of Welfare

DANIEL T. SLESNICK[1]

1. *Introduction*

THE MEASUREMENT of welfare forms the foundation of public policy analysis. A full consideration of taxes, subsidies, transfer programs, health care reform, regulation, environmental policy, the social security system, and educational reform must ultimately address the question of how these policies affect the well-being of individuals. While centrally important to many problems of economic analysis, confusion persists concerning the relationship between commonly used welfare indicators and well-established theoretical formulations. For years it seemed that theoretical and empirical research developed along parallel lines with little, if any, cross-fertilization.

The gap between theory and applications in welfare economics is closing. In the first part of the survey, I examine the conceptual and empirical issues related to welfare measurement at the micro level. Over the years, analysts have relied heavily on consumer's surplus to measure the welfare effects of changes in prices and incomes. This largely reflects the modest data requirements and the ease with which it is implemented. Despite its ubiquity, it is now widely accepted that consumer's surplus should not be used as a welfare measure, although there is less agreement as to why. The critical issues of the debate are summarized in Section 2.1.

The last twenty-five years have seen great progress in the development of measures that are ordinally equivalent to household utility. These welfare indicators, which are described Section 2.2, do not have the same limitations as consumer's surplus, but are easy to implement and have the same data requirements. Each infers changes in welfare from the consumption behavior of households, using either econometric methods or index numbers. As a result, there is substantial overlap between the empirical issues associated with demand modeling (both static and intertemporal) and the general problem of welfare measurement. I summarize several important econometric issues in Section 2.3, and recent developments using the index number framework are presented in Section 2.4. The first part of the survey concludes with a description of three applications that illustrate many of the general empirical issues.

While welfare measurement at the micro level is of independent interest, of greater concern is the well-being of groups of households. The second half

[1] University of Texas, Austin. I would like to thank Don Fullerton, Dan Hamermesh, Arthur Lewbel, Bob Pollak, Max Stinchcombe, and Frank Wolak for providing thoughtful comments and suggestions on previous drafts of this paper. Any remaining errors are solely my responsibility.

of the survey will examine the aggregation problem in applied welfare economics. The most common approach is to assume the existence of a representative consumer and use market demands as the basis for the measurement of social welfare. This is unappealing both because distributional issues are ignored and because much evidence shows that aggregate demands are inconsistent with the behavior of a single representative agent.

To go beyond this framework requires normative judgements concerning the measurability and comparability of welfare across heterogeneous agents. In most applications, these assumptions are implicit but no less arbitrary than those based on explicit social welfare functions. For example, the common practice of representing social welfare by per capita income implicitly assumes cardinally measurable welfare that is fully comparable across households. Since incomes are simply averaged across the population, the implicit social welfare function is utilitarian and the marginal utility of income is assumed to be constant.

Social choice theory is a natural framework for examining the normative basis for aggregate welfare measures used in practice. How does one get around Arrow's Impossibility Theorem? If this is feasible, what additional assumptions are required? Beyond the many conceptual issues, what types of household welfare functions and data are required to "build" a function that can be used to measure aggregate welfare? Is the aggregation problem an arcane issue of interest to only a handful of social choice theorists? In fact, the measurement of group welfare lies at the heart of some of the most pressing issues in applied economics. In some cases, the method of aggregation has little effect on our conclusions, while in others it is vitally important.

As in all surveys of finite length, some noteworthy contributions are ignored here or given only cursory consideration. I intend to restrict the focus to developments in the "welfarist" tradition in which well-being is derived from the consumption of goods and services. This seems to be the natural focus since most empirical work is in this framework, although limiting the scope has the unfortunate consequence of short-changing several important developments in welfare economics. For example, Sen's capabilities approach to welfare measurement is not considered in depth largely because of the relative scarcity of empirical applications. The survey-based or "subjective" approach pioneered by van Praag, Kapteyn and others is fundamentally different from welfare estimates based on households' revealed preferences and is also excluded. A growing body of theoretical and empirical work suggests that allocations of goods within households have important consequences for welfare measurement. While this issue will be discussed, it will not play a prominent role in this survey. Finally, let me emphasize that no attempt is made to survey the voluminous social choice literature. The intention is to use key results on the possibility of consistent social choice to describe a framework for examining alternative approaches to the measurement of aggregate welfare.

2. *Measuring Household Welfare*

It is natural to organize the discussion of developments in welfare economics by the level of aggregation at which they are applied. I begin with the issue of welfare measurement for households and focus on the welfare effects of price and income changes.

Infinite divisibility is assumed and welfare is broadly defined as the money needed to maintain a constant level of utility.

2.1 *What's Wrong with Consumer's Surplus*

It has now been over twenty-five years since Arnold Harberger (1971) published his open letter to the profession in which he proposed a set of guidelines for applied welfare economists. Simply put, consumer's surplus should be used to measure individual welfare, and social welfare should be based on its unweighted sum over the population. A casual examination of the large number of empirical cost-benefit analyses suggests that many practitioners have taken Harberger's advice to heart. Consumer's surplus is the overwhelming choice as a welfare indicator, largely, I would guess, because it is an intuitive concept with modest data requirements.

From its inception over 150 years ago, the validity of consumer's surplus as a welfare measure has been questioned.[2] Harberger's proposal generated a new round of debate that evoked many of the same arguments that had circulated for years. The discussion took a more productive turn because analysts not only demonstrated what was wrong with consumer's surplus but suggested reasonable alternatives for actually "doing" welfare economics.

As a point of departure, I summarize the conceptual problems with consumer's surplus. Chipman and Moore (1976b, 1980) have shown that for a wide variety of applications, consumer's surplus is a valid measure only under restrictive conditions on preferences. Define M_k to be the level of expenditure of the kth household,[3] $p = (p_1, p_2, \ldots, p_n)$ is a vector of prices, A_k is a vector of demographic characteristics and $x_i(p,M_k,A_k)$ is the demand function for good i. In the simplest case of a change in the price of a single good (say, commodity 1) from p_1^0 to p_1^1, the change in consumer's surplus is the negative of the change in the area under the demand curve:

$$\Delta CS_k = -\int_{p_1^0}^{p_1^1} x_1(t, p_2, \cdots, p_n, M_k, A_k)dt. \quad (1)$$

Is consumer's surplus an exact representation of the change in welfare? That is, is it single valued and ordinally equivalent to the change in utility? A necessary condition is that demands are generated by a rational consumer who maximizes utility subject the constraint of limited resources. Unless consumers have optimized and are at the boundaries of their budget sets, it is impossible to assess the welfare implications of changes in prices and incomes. The maintained hypothesis must therefore include the condition that demands are "integrable" and consistent with a well-behaved utility function.[4]

Under these conditions, an indirect utility function $V(p, M_k, A_k)$ represents the maximum attainable utility at prices p and expenditure M_k. Application of Roy's

[2] An historical perspective on the early debate over the use of consumer's surplus is provided by John Chipman and James Moore (1976b).

[3] While income is commonly used as an argument of the demand function, in a static context the appropriate variable is total expenditure. Moreover, I'll take the basic observational unit to be a household rather than an individual since this is the way expenditure data are usually reported. Of ultimate concern, of course, is the welfare of the individuals who belong to the household, which is an issue that will be considered further when I examine the intrahousehold allocation of resources.

[4] The integrability conditions are described by Leonid Hurwicz and Hirofumi Uzawa (1971). If demands are integrable and generated by a well-behaved utility function they must be nonnegative, summable, homogeneous of degree zero in prices and total expenditure, and the Slutsky matrix of compensated price effects must be symmetric and negative semidefinite.

Identity provides the link between the demands and the indirect utility function:

$$x_1(p, M_k, A_k) = -\frac{\partial V(p, M_k, A_k)/\partial p_1}{\partial V(p, M_k, A_k)/\partial M_k}. \quad (2)$$

If the marginal utility of income is constant, substitution of (2) into (1) yields an explicit expression that is ordinally equivalent to the change in utility:[5]

$$\Delta CS_k = \frac{V(p^1, M_k, A_k) - V(p^0, M_k, A_k)}{\partial V/\partial M_k}.$$

For a single price change, constancy of the marginal utility of income is sufficient to ensure that consumer's surplus provides an exact measure of the change in welfare.

While this condition is restrictive, the problems become more serious with changes in more than one price. The change in consumer's surplus must be evaluated using a line integral defined over the path of price changes:

$$\Delta CS_k = -\int_{p^0}^{p^1} \Sigma x_i(p, M_k, A_k) dp_i, \quad (3)$$

where p^0 and p^1 are the initial and final price vectors.

This extension is conceptually straightforward but suffers from the problem that line integrals generally depend on the paths over which they are evaluated. Since we only observe the initial and final price vectors, the change in consumer's surplus (3) must be independent of the path that, in turn, occurs if the uncompensated price effects are symmetric:[6]

$$\frac{\partial x_i}{\partial p_j} = \frac{\partial x_j}{\partial p_i} \qquad \text{for } all\ i \neq j.$$

A rational consumer will exhibit compensated price effects that are symmetric, but uncompensated demands will have this property only if the income effects are always equal. This requires equality of the income elasticities which can only occur if they are equal to one and preferences are homothetic.[7] This is a serious problem from a practitioner's perspective since now almost 150 years of empirical evidence demonstrates that demand patterns are inconsistent with homotheticity.[8]

If we want to measure welfare in the most general circumstance, in which prices and total expenditure change, we must have path independence over both sets of variables:

$$\Delta CS_k = -\int_Z \Sigma x_i(p, M_k, A_k) dp_i + (M^1 - M^0), \quad (4)$$

where Z is a path between (p^0, M^0) and (p^1, M^1). Chipman and Moore (1976b) have shown there are no circumstances in which (4) is path independent and ordinally equivalent to the the change in welfare for a well-behaved utility function.[9]

John Hicks' (1942) approach to welfare measurement does not have the same shortcomings as consumer's surplus. When both prices and total expenditure change, the Hicksian surplus

[5] The reliance of consumer's surplus on the assumption of constant marginal utility of income is discussed by Chipman and Moore (1976b) among many others.

[6] See, for example, Angus Taylor and Robert Mann (1972), pp. 500–504.

[7] This follows from "Engel aggregation" which requires the share weighted sum of the income elasticities to be equal to one. The requirement of homotheticity for path independence has been shown by Eugene Silberberg (1972) and Chipman and Moore (1976b).

[8] See the survey by Hendrik Houthakker (1957) marking the centennial anniversary of Engel's Law.

[9] Chipman and Moore (1976b) also consider the case in which income and all prices except one (taken to be the numeraire) change. In this case, consumer's surplus provides an exact measure of the change in welfare if preferences are quasi-linear in the numeraire. This implies that the demands for all goods other than the numeraire are independent of income which, again, is an empirically untenable restriction.

measure is exactly analogous to (4) once we substitute compensated for uncompensated demand functions:

$$\Delta HS_k = -\int_Z \Sigma x_i^c(p, V, A_k)dp_i + (M^1 - M^0), \quad (5)$$

where $x_i^c(p, V, A_k)$ is the compensated demand for the ith good evaluated at utility level V. Compensated price effects are symmetric, so the line integral in (5) is path independent and the surplus measure is single-valued.

For simple binary comparisons, the utility level at which (5) is evaluated is often treated as a matter of little importance. If (5) is calculated at the utility attained at prices p^1 and expenditure M^1 (denoted V^1), we obtain a generalized version of the equivalent variation:

$$EV_k = M(p^0, V^1, A_k) - M(p^1, V^1, A_k) + (M^1 - M^0)$$
$$= M(p^0, V^1, A_k) - M(p^0, V^0, A_k), \quad (6)$$

where $M(p, V, A_k)$ is the expenditure function. Note that if only prices change and there is no change in total expenditure, (6) simplifies to the conventional representation of the equivalent variation. Not only is the generalized equivalent variation (also referred to as the equivalent gain by Mervyn King 1983) single valued, but it is ordinally equivalent to the change in utility (i.e. it is positive if and only if $V^1 > V^0$) .

The utility level at which (5) is evaluated is important for multiple comparisons of price-expenditure changes. Such comparisons are the rule rather than the exception because it is often the case that several policies are compared with the status quo or some other benchmark. The equivalent variation (6) will give an ordering of outcomes that is identical to that based on utility levels. If, however, (5) is calculated at $V^0 = V(p^0, M^0, A_k)$, we obtain the generalized compensating variation:

$$CV_k = M(p^0, V^0, A_k) - M(p^1, V^0, A_k) + (M^1 - M^0)$$
$$= M(p^1, V^1, A_k) - M(p^1, V^0, A_k). \quad (7)$$

Because the utility levels are "cardinalized" using different prices for each set of binary comparisons, the ordering of multiple price-expenditure outcomes using (7) need not match the ordering based on utility levels. Chipman and Moore (1980) have shown that consistent rankings of outcomes require restrictions on preferences that are the same as those needed to resurrect consumer's surplus.

The generalized equivalent variation (6) resolves the conceptual problem of welfare measurement, but it does not help the practitioner because (it was originally thought) compensated demands and, therefore, the Hicksian surplus measures are unobservable. Robert Willig (1976) made the first attempt to tackle this problem by demonstrating that, in the case of a single price change, (observable) estimates of consumer's surplus can be used to provide a good approximation to the equivalent or compensating variation. Since these welfare indicators provide upper and lower bounds for consumer's surplus, the accuracy of the approximation depends on the magnitude of the income effect.

Willig gave a number of examples illustrating the accuracy of his approach. It was only a matter of time, though, before counterexamples involving goods with large income effects began to appear. Jerry Hausman (1981) considered the problem of labor supply and found Willig's method to be inaccurate.[10] It is

[10] While his basic point was correct, Hausman's calculation of the error was overstated by an arithmetic error. See Robert Haveman, Mary Gabay, and James Andreoni (1987). Hausman also pointed

worth reemphasizing, however, that with multiple price and expenditure changes, consumer's surplus is not single-valued and is of no use in providing an approximation to the Hicksian surplus measures.[11]

2.2 *Exact Measures of Individual Welfare*

If consumer's surplus cannot be used as an approximation, is there any hope of measuring the equivalent or compensating variations? Immediately following the publication of Willig's paper, empirical procedures were developed that had the same data requirements as consumer's surplus without the conceptual problems. Each method begins with assumed forms for the vector of demand functions $x(p, M_k, A_k)$. As long as the integrability conditions are satisfied, they contain all of the information necessary to recover the underlying utility and expenditure functions. The recovery procedure involves "integrating backwards" from the demands which, in practice, can be difficult if the functions are the least bit complicated. The complexity diminishes significantly if we restrict consideration to changes in the price of a single good and assume that demands are linear in prices and expenditure.

To illustrate, assume only two goods; the commodity of interest (denoted good 1) and a composite commodity with a price normalized to one. By Roy's Identity, the demand for the good is related to the indirect utility function by:

$$x_1(p_1, M_k, A_k) = -\frac{\partial V(p_1, M_k, A_k)/\partial p_1}{\partial V(p_1, M_k, A_k)/\partial M_k}. \quad (8)$$

As long as the demands are consistent with utility maximization, equation (8) is a partial differential equation that can be solved to obtain the indirect utility function and corresponding expenditure function.

Hausman (1981) described an analytic solution to an equation of this type when the demand function is linear or log-linear in price and expenditure. Suppose the demand function is:

$$x_1 = \gamma_p p_1 + \gamma_m M_k + \gamma_A A_k,$$

where γ_p, γ_m, and γ_A are unknown parameters that can be estimated econometrically. Equation (8) can be solved to yield an expenditure function of the form:

$$M(p_1, V, A_k) = Ve^{\gamma_m p_1} - (1/\gamma_m)[\gamma_p p_1 + (\gamma_p/\gamma_m) + \gamma_A A_k]. \quad (9)$$

Given the expenditure function, an exact measure of the change in welfare is easily calculated, and approximations are unnecessary. Note also that only linear regression methods are required to estimate the unknown parameters, and the data requirements are identical to those needed to compute consumer's surplus.

This simplicity does not come without costs. Closed form solutions to the partial differential equation shown in (8) can be obtained for only a limited class of demand functions. Hausman provides solutions to the linear and log-linear specifications but also showed that the problem becomes much more complicated when quadratic terms in expenditure are added. For ease of implementation, he purposely restricted consideration to a partial equilibrium framework in which only one equation is estimated and the welfare effects of a single price change are evaluated. His approach is of little use in the general circumstance in which a vector of prices and expenditure change because the expenditure function cannot be recovered from a single equation.

---

out that, even if Willig's method provides a decent approximation of the change in utility, it is not a good approximation of the deadweight loss.

[11] This point was illustrated by George McKenzie (1979).

More general approaches have been developed along several different lines. The problem of recovering the underlying utility function from a system of demand equations can be solved by beginning with an assumed form of the indirect utility function. Application of Roy's Identity yields a set of demand equations that can be estimated using data on demands, prices, and demographic characteristics. Since the form of the utility function is assumed from the outset, it is unnecessary to solve a complex system of partial differential equations. In one of the first applications of this approach, John Muellbauer (1974) assumed that demands were consistent with the Stone-Geary utility function and estimated the unknown parameters by fitting the linear expenditure system to household budget data.[12] Given the parameter estimates, the expenditure function was obtained directly and used to measure changes in welfare that resulted from multiple price and expenditure changes.

To illustrate further, consider the utility function proposed by Dale Jorgenson, Lawrence Lau and Thomas Stoker (1982):

$$\ln V(p,\ M_k, A_k) = \alpha_p{}'\ln p + 1/2\ln p'B_{pp}\ln p - D(p)\ln M_k + \ln p'B_{pA}A_k, \quad (10)$$

where $D(p) = -1 + \iota B_{pp}{}'\ln p$ and α_p, B_{pp} and B_{pA} are unknown parameters. These can be estimated by applying the logarithmic version of Roy's Identity to (10) to obtain a system of share equations:

$$w_k = \frac{1}{D(p)}(\alpha_p + B_{pp}\ln p - \iota B_{pp}\ln M_k + B_{pA}A_k), \quad (11)$$

where w_k is the vector of budget shares of all goods consumed by the kth household. Using the estimated parameters, welfare changes can be calculated using the implied expenditure function:

$$\ln M(p, V, A_k) = \frac{1}{D(p)}(\alpha_p{}'\ln p + 1/2\ln p'B_{pp}\ln p - \ln V + \ln p'B_{pA}A_k).$$

While this is more general than Hausman's approach, it has disadvantages. For an assumed form of the utility function, the functional forms of the demands are the same for every good, which may hinder the ability of the model to fit the data. Moreover, Hausman argues that sound econometric practice should begin with the "best" specification of the function to be fit to the data, which, in this application, is the set of demand equations.

Is it possible to start with an arbitrary demand system (rather than a utility function) and measure the welfare effects of multiple price changes? Two elegant procedures have been proposed that require more complicated calculations but do not impose restrictions on the form of the demand functions other than the standard integrability conditions. The first method is based on an approximation to the indirect money metric utility function, which is defined as:

$$\mu(p, M_k, A_k; p^0) = M(p^0, V(p, M_k, A_k), A_k).$$

Developed originally by Lionel McKenzie (1957), the indirect money metric utility function serves as an exact representation of utility. In fact, the generalized version of the equivalent variation defined in (6) is simply the difference between the indirect money-metric utility functions evaluated at the initial and final price and expenditure vectors.

For an arbitrary, unknown function μ, McKenzie and Ivor Pierce (1976) showed that monetary measures of the change in utility can be measured using a Taylor's series expansion about the initial equilibrium:

[12] Another early example of this approach is Harvey Rosen (1978).

$$\Delta\mu = \frac{\partial\mu}{\partial p}'\Delta p + (1/2)\Delta p'\frac{\partial^2\mu}{\partial p\partial p'}\Delta p + \left(\frac{\partial\mu}{\partial M} + \frac{\partial^2\mu}{\partial p\partial M}'\Delta p + 1/2\frac{\partial^2\mu}{\partial M^2}\Delta M\right)\Delta M + R, \quad (12)$$

where R represents higher order terms in the series. At first glance, (12) does not appear to be useful for empirical applications because the derivatives of μ cannot be observed.

McKenzie and Pearce noted that money-metric utility is identical to total expenditure when evaluated at the reference prices. The marginal utility of income is therefore equal to one, and all higher order income derivatives are zero. This fact can be used with Roy's Identity to represent (12) as a function of the Marshallian demands and their income and price derivatives:

$$\Delta\mu = -x'\Delta p - (1/2)\Delta p'\left(\frac{\partial x}{\partial p} - x\frac{\partial x}{\partial M}'\right)\Delta p + \left(1 - \frac{\partial x}{\partial M}'\Delta p\right)\Delta M + R. \quad (13)$$

Given knowledge of the demand functions and the magnitudes of the price and expenditure changes, one has all of the information necessary to get as accurate an estimate of the change in utility as desired. As a practical matter, though, the analyst has no information as to the number of terms in the Taylor's series that are required to obtain an accurate estimate. If higher order terms are required, the calculations become more cumbersome (but not intractable).

Yrjo Vartia (1983) developed an algorithm that estimates the expenditure function numerically to any desired level of accuracy. Let $p(t)$ and $M_k(t)$ describe the paths of price and expenditure changes for $0 \leq t \leq 1$. Using Roy's Identity, he shows that as prices and expenditure change, the movements of demands along an indifference curve are described implicitly by the first order differential equation:

$$\frac{dM_k(t)}{dt} = \sum_{i=1}^{n} x_i(p(t), M_k(t), A_k)\,\frac{dp_i(t)}{dt}.$$

Integrating over t yields an equation that can, in principle, be solved for $M(p(t), V^0, A_k)$ which is the centerpiece in the Hicksian surplus calculations:

$$M(p(t), V^0, A_k) - M(p^0, V^0, A_k) = \sum_{i=1}^{n}\int_0^t x_i(p(t), M(p(t), V^0, A_k), A_k)\,\frac{dp_i(t)}{dt}\,dt. \quad (14)$$

Vartia described several algorithms that can be used to solve this equation numerically over a path $p(t)$ so that, when evaluated at $t = 1$, we obtain $M(p^1, V^0, A_k)$. As long as the demands satisfy the integrability conditions, the solution to (14) will be independent of the price path used in the algorithm. This method is valid for multiple price and expenditure changes and, because a closed-form solution is unnecessary, facilitates flexibility in estimating demand patterns. This is an important advantage although it is worth noting that multiple comparisons of welfare changes require repeated application of the algorithm.This is more inconvenient than evaluating an explicit form of the expenditure function although, given the current computational power that is widely available, this is less of an issue today than it was when Vartia originally proposed his approach.

2.3 *Empirical Issues*

The conceptual problem of how to measure household welfare using ordinary Marshallian demand functions has largely been resolved. Since these functions are unknown, econometric methods are used to estimate them statistically. Welfare measurement is, as a result, indelibly linked to empirical demand analysis and issues related to the specification, estimation, and data requirements of the model must be addressed. Comprehensive surveys of the voluminous literature on demand modeling are presented by Angus Deaton and Muellbauer (1980a), Richard Blundell (1988), Robert Pollak and Terence Wales (1992), and Arthur Lewbel (forthcoming).

The standard approach is to specify a functional form for the demand function and estimate the unknown parameters using regression methods. But what function should be chosen a priori? A misspecification will lead to biased estimates and erroneous inferences concerning the magnitudes and, possibly, even the directions of the changes in well-being. It is important, therefore, to develop a framework with enough flexibility to allow the data to describe accurately the relationships between demands and prices and total expenditure.

The specification problem is not limited to the choice of functional form. Households consume many goods, and data limitations prevent the estimation of every demand function. The model can be simplified either by restricting consideration to a subset of the goods or aggregating all items into broad commodity groups. In many applications this is not a matter of choice but of necessity because the data have information on only a small number of categories. In fact, it may be the case that we have data on only a single commodity such as labor supply. Under what conditions is it possible to analyze a subset of all goods in isolation from the others? If these conditions hold, what can be learned about changes in welfare from the subsystem of demand functions?

Even at the micro level, empirical implementation requires the resolution of a potentially serious aggregation problem. Expenditure data are reported for entire households by simply adding up the spending of all members. While we have the information at the household level, it is the welfare of the individual members that we are concerned about. This requires inferences concerning the intrahousehold distribution of consumption and the evaluation of welfare derived from these goods for each member. As we shall see, this research has raised important questions that have proved difficult to answer and is likely to be a fruitful area for future work.

Functional Form Specification.[13] Until the recent development of nonparametric statistical models, empirical demand analysis began with the specification of the forms of the demand functions (or the corresponding utility function) to be fit to the data. Perhaps because of computational limitations, the earliest applications used specifications that could be estimated using linear or only slightly nonlinear regression methods. In his pioneering study, Richard Stone (1954) developed and implemented the linear expenditure system and, much later, Hausman (1981) restricted attention to functional forms that could be estimated using linear regression. While easy to implement, linear demand functions may not be sufficiently flexible to measure demand responses to price and expenditure changes. The estimated elasticities may reflect the functional form assumption rather than the demand patterns revealed by the data.

Out of this concern grew the development of what have become known as "flexible functional forms". Any specification is almost surely incorrect, and the best one can hope for is an approximation to the demand or utility function.

[13] This subsection has benefited greatly from comments and suggestions from Arthur Lewbel.

The indirect utility function $V(p_k^*)$ provides a local, second-order approximation to the true function $V^*(p_k^*)$ about the point $\overline{p}_k^*$ if

$$V^*(\overline{p}_k^*) = V(\overline{p}_k^*)$$
$$\frac{\partial V^*(\overline{p}_k^*)}{\partial p_k^*} = \frac{\partial V(\overline{p}_k^*)}{\partial p_k^*}$$
$$\frac{\partial^2 V^*(\overline{p}_k^*)}{\partial p_k^* \partial p_k^{*\prime}} = \frac{\partial^2 V(\overline{p}_k^*)}{\partial p_k^* \partial p_k^{*\prime}},$$

where $p_k^* = p/M_k$ is the vector of normalized prices. Erwin Diewert (1974) defined functions V that satisfy these conditions to be (second-order) flexible functional forms.

The translog utility function developed by Laurits Christensen, Jorgenson and Lau (1975) is an example of a model that satisfies the condition of local second-order flexibility (a special case is shown in equation 10). A critical feature is that enough parameters are used to ensure that the demand elasticities are determined, without prior constraints, by the data. Other second-order flexible functional forms include the generalized Leontief form developed by Diewert (1971), William Barnett and Yul Lee's (1985) minflex Laurent demand system, and the AIDS model developed by Deaton and Muellbauer (1980b).

As mentioned previously, the recovery of information on preferences from demands requires that they be consistent with a rational consumer who maximizes utility subject to a budget constraint. This requirement imposes cross-equation restrictions on the demands that limit their ability to fit the "true" underlying functions. In the context of welfare measurement, the challenge is to maintain the flexibility of the demands functions while imposing the restrictions necessary to ensure the existence of a well-behaved indirect utility function.

If demands are consistent with utility maximization, the spending on all goods must sum to total expenditure M_k (summability) which makes one of the equations superfluous in the estimation procedure. Monotonicity of the utility function implies that the demand for each good must be nonnegative at all prices and expenditure levels (nonnegativity). This is typically difficult to impose at all data points, and, for example, the translog model shown in (11) does not satisfy this restriction globally. The condition that demands be homogeneous of degree zero in prices and expenditures (homogeneity) imposes further constraints on the price and expenditure parameters within each equation. Symmetry of the Slutsky matrix of compensated price effects (symmetry) imposes cross-equation restrictions on the price parameters, while quasiconvexity of the utility function requires it to be negative semidefinite. The latter condition imposes complicated nonlinear inequality restrictions on the demand parameters to ensure that the eigenvalues of the Slutsky matrix are nonpositive. Taken together, the full complement of restrictions implied by consumer rationality are quite strong and can limit the ability of functional forms to fit the data well.[14]

Recent work has focused on flexible demand models that are linear in functions of total expenditure:

$$x_i(p, M_k) = \sum_{l=1}^{L} h_{il}(p) f_l(M_k), \qquad (15)$$

where, for notational simplicity, I have suppressed the effects of demographic

[14] In particular, it is possible that the constraints reduce the number of free parameters to such an extent that it is impossible to provide a second order approximation to an arbitrary function. The relationship between flexibility and the integrability conditions has been considered by Diewert and Wales (1987) in the production context.

characteristics on the demand functions. This class of models is particularly important because the micro level demands aggregate exactly to yield macro demands that are functions of statistics that depend on the level and distribution of total expenditure.[15] Gorman (1981) showed that if demands have this linearity property and are consistent with utility maximization, then the maximum rank of the $(n \times L)$ matrix $H(p)$ formed from the elements $h_{il}(p)$ is three. Of special interest are the models that Lewbel (1990) describes as "full rank" demand systems where the rank of $H(p)$ is equal to L. In these models, the econometrician achieves the greatest flexibility in describing the relationship between demands and total expenditure with the tightest parametric specification. Lewbel characterizes all full rank demand models of the form shown in (15) and finds that many of the commonly used (aggregable) demand systems are of this type.[16]

The importance of Gorman's and Lewbel's results is that they suggest that many of the demand systems used in applications are not as general as they could be. The translog demand system (11) and the AIDS model are full rank demands systems that depend on two functions of total expenditure. Both are functions of M_k and $M_k \ln M_k$ and belong to the general class of models referred to as "price independent generalized log linear" (PIGLOG) by Muellbauer (1976). Rank two models such as these may not adequately describe the relationship between the demands and total expenditure.

In fact, recent evidence presented by Lewbel (1991) and Blundell, Panos Pashardes and Guglielmo Weber (1993) based on Engel curves indicates that rank three extensions of the PIGLOG model provide better descriptions of expenditure patterns in both the U.S. and the United Kingdom. In their models, demands depend on three functions of total expenditure: M_k, $M_k \ln M_k$ and $M_k(\ln M_k)^2$.[17] Not only does the rank three extension provide a better description of the data, but it demonstrates that if one restricts consideration to exactly aggregable models, higher order approximations that yield more than three functions of expenditure are superfluous. More important for our purposes, James Banks, Blundell and Lewbel (1995) show that a rank three generalization of the AIDS model leads to different conclusions concerning the welfare effects of commodity tax reform relative to the standard rank two AIDS model.

While the development of flexible functional forms represented an important step forward in empirical demand modeling, several practical problems emerged. Halbert White (1980) pointed out that the estimated function might not provide an accurate approximation to the true underlying utility function even in large samples. Part of the problem is that the point of approximation is unknown and little can be said about the behavior of the estimated function at other data points. Ignoring the remainder term of the series, moreover, results in a misspecification that leads

[15] For further discussion of exactly aggregable demand models, see Lau (1982), William Gorman (1981), and the survey by Stoker (1993).

[16] Lewbel (1989c, 1991) has extended these results for demand specifications that are more general than (15). For example, if f_l is a function of "deflated expenditure" rather than expenditure alone, the maximum possible rank of the demand system increases to four.

[17] Lewbel (1990) has shown that the other possible forms of the full rank three models are "generalized quadratic" and demands that depend on trigonometric functions. In the former case, the three functions are M_k, $M_k^{1-\alpha}$ and $M_k^{1+\alpha}$ for arbitrary values of α.

to biased and inconsistent estimates. Beyond these statistical concerns, the quality of the approximation provided by a flexible functional form depends on the behavior of the remainder term. There is no reason a priori to expect the truncation error to be small even around the point of expansion. Globally, very little can be said.[18]

Some researchers have addressed these issues by using semiparametric or nonparametric methods to model demands. Ronald Gallant (1981) and Ibrahim Elbadawi, Gallant and Geraldo Souza (1983) proposed a hybrid of the parametric and nonparametric techniques by appending a Fourier series expansion of V to a second order flexible functional form such as the translog:

$$V(p_k^*) = \alpha' \ln p_k^* + 1/2 \ln p_k^{*\prime} B \ln p_k^* + \sum_{a=1}^{A} F_a(\lambda_a' p_k^*).$$

In this model, F_a is a function of sines and cosines representing the elements of the Fourier expansion and λ_a is a multidimensional index. The model is parametric in the usual sense but nonparametric because the number of terms A included in the Fourier series depends on the sample size. As the sample size grows, the semiparametric model is a sequence of successively more complicated functional forms that, in the limit, provide a global approximation to an arbitrary utility function and all of its derivatives. In this application, the derivatives are particularly important because they represent the demands (the first derivatives) and the elasticities (the second derivatives).

The semiparametric approach is attractive because the advantages of the parametric models are retained. The integrability conditions are represented as parameter restrictions and the entire system of demand equations can be estimated using the usual multivariate statistical methods. Moreover, an explicit representation of the underlying indirect utility function is obtained that can be used to estimate the welfare effects of price and expenditure changes. Michael Creel (1997) provides an example of welfare calculations that are based on the Fourier model.

These methods, however, have their own problems. The approximation properties are valid only with infinite samples which, of course, are never found in practice. With finite samples, the semiparametric model is essentially equivalent to a parametric model, and we have no reason to expect it to perform better than any other. For example, even with appropriately-scaled data, it is not obvious that the trigonometric components of the truncated Fourier system would accurately describe consumers' responses to price and expenditure changes. While cross-validation and other diagnostics are being used with greater frequency, the bottom line is that the magnitude of the truncation error induced by finite samples is difficult to assess. As a result, a conflict persists between overparameterization (as more terms of the series are added) and the inaccuracy that results from truncating the series.

There have been several recent efforts to estimate demands using a fully-nonparametric specification.[19] Assume the system is given by:

[18] White (1980) gives several examples in which the biases are large. R.P. Byron and Anil Bera (1983) note, however, that White's results concerning the inaccuracy of the regression function is overstated because he considers only a first order approximation. They demonstrate that second order approximations do much better even though the bias persists.

[19] See Wolfgang Hardle (1990) for an introduction to these nonparametric statistical methods.

$$x_k = x(p, M_k, A_k).$$

The demand function for the ith good $x_i(.)$ is estimated using a kernel regression of the form:

$$\hat{x}_i(.) = \frac{(1/K)\sum_{k=1}^{K} x_{ik}G(.)}{(1/K)\sum_{k=1}^{K} G(.)} \qquad (i = 1, 2, \ldots, n),$$

where $G(.)$ is the kernel used to estimate the joint density of the regressors.

Nonparametric methods are appealing because they avoid the vexing issue of functional form misspecification and its contaminating effects on statistical inference. Recent empirical work in demand analysis using variants of these techniques include Herman Bierens and Hettie Pott-Buter (1990), Banks, Blundell, and Lewbel (1995), Hardle, Werner Hildenbrand and Michael Jerison (1991), Hausman and Whitney Newey (1995) and Lewbel (1991, 1995). This is a promising new approach, but it introduces additional problems. While an explicit functional form for the demand function is unnecessary, it requires choices for the appropriate kernel $G(.)$ and the bandwidth to be used in the smoothing. Both choices can have a substantial influence on the estimation results. Both nonparametric and semiparametric methods require large numbers of observations with substantial variation in the regressors. While household budget surveys have enough observations, they typically have little or no price information, which often precludes the use of these techniques. Most applications have focused on the relationship between demand and total expenditure, and they have ignored price effects.

Another impediment to the application of these estimators is the "curse of dimensionality" in which the speed of convergence of the estimators to the true values is inversely related to the number of regressors. Out of this concern, demand applications typically use only one or two explanatory variables and have modeled the demand using a single equation. As we shall see below, this requires separability assumptions, which likely result in a misspecification of the model. This form of specification error must be weighed against the advantages of the nonparametric estimator in discerning the higher order derivatives of the functional form of the demand function.[20] Analyzing a single equation also avoids the issue of cross equation restrictions implied by the theory of consumer behavior. It is unclear how tractable or practically useful nonparametric methods are in the context of modeling a complete, multivariate demand system with the imposition of the full set of integrability restrictions.[21]

While nonparametric methods are not uncommon in the analysis of Engel curves, relatively few studies have used these techniques to draw inferences concerning changes in welfare. Hausman and Newey (1995) exploited cross sectional price variation to obtain a nonparametric estimator of a single demand function $\hat{x}(p, M_k)$ (gasoline), and solved numerically for the underlying expenditure function using Vartia's method described in (14). They analyzed the welfare effects of gasoline taxes and, after all was said and done, found little difference between the nonparametric estimates and those obtained using a simple linear parametric

[20] I thank Arthur Lewbel for this point.

[21] Some progress has been made on this front. Hausman and Newey (1995) and Lewbel (1995) describe nonparametric tests of the integrability conditions although the tests are quite cumbersome and difficult to implement.

model. Deaton (1989) summarized the welfare effects of differential price changes by a single variable, and used kernel regression methods to examine how they changed for households with different incomes.

Separability over Goods and Time. How many demand equations should be used to measure welfare? Individuals consume literally hundreds of goods (thousands if one distinguishes goods by the date of consumption) so that, with the typical constraints imposed by limited sample sizes, there is no hope of estimating a complete demand system with any degree of precision. Parsimonious specifications can be achieved through some form of aggregation over commodities, which is justified by several layers of assumptions of separability of preferences over time and within periods.

The welfare measures described in the previous section are based on a single period framework in which the effects of changes in the prices and expenditures in other periods are ignored. The welfare effect within a period is independent of other time periods if there is intertemporal separability. This is usually assumed to take the form of two-stage budgeting, in which the household decides the total amount that will be consumed in each period and then, conditional on this choice, the allocation across goods within a period. In this context, it is appropriate to measure the single-period welfare effect using only the prices and expenditures observed in that time period.[22]

The two-stage budgeting framework has several implications for the static welfare indicators such as (6) and (7). First, it influences the functional structure of both the within-period and intertemporal utility functions. Gorman (1959) has shown that spending patterns are consistent with two-stage budgeting if either the lifetime utility function is homothetically separable or the top-level utility function is additive and the within-period functions are of the Gorman polar form.[23] Since homothetic preferences have no empirical support, additivity is the more palatable assumption. Note that even in the absence of two-stage budgeting, the intertemporal separability assumption alone precludes the inclusion of lagged values of the demands as arguments of the demand functions. This is restrictive because dynamic specifications of this type have been found to be empirically important in explaining consumption behavior in the context of habit formation, addiction and durability models.[24]

A second implication of two-stage budgeting is that within-period demands must be functions of total expenditure (or full expenditure if labor supply is endogenous) rather than income. Total expenditure (along with the marginal utility of wealth) serves as a "sufficient statistic" that links the within-period model to the intertemporal allocation of consumption. The use of income in static demand models, which is very common, results in a potentially serious misspecification because of its systematic understatement of expenditure for low-income households and the

[22] A separate question is how one should measure the intertemporal change in well-being. If the price and expenditure changes cause a rearrangement of consumption over time, the "snapshot" welfare indicators alone will give a biased picture of the change in lifetime utility. This point will be elaborated later when the model is extended to measure changes in intertemporal welfare in Section 2.5.

[23] For further discussion, see Charles Blackorby, Dan Primont, and Robert Russell (1978).

[24] As examples, see Gordon Anderson and Blundell (1982), Gary Becker, Michael Grossman, and Kevin Murphy (1994) and Costas Meghir and Weber (1996) among others.

reverse for those with high incomes. This turns out to be particularly important when welfare measures are used to examine distributional issues such as inequality or poverty.

Given intertemporal separability, commodity aggregation for the within-period demands is an additional problem. The analyst often has information on only a subset of the demands; sometimes a single commodity. In measuring the welfare effects of a change in wages, for example, it may be that only labor supply data are available but comprehensive consumption data are not. This is the situation if one is using, say, the *Current Population Surveys* in the United States. Under what conditions can we measure the within-period change in welfare from a single demand equation?

In general, what is required is the aggregation of consumption into a composite commodity. One way to do this is to invoke the Hicks-Leontief composite commodity theorem, which states that a group of goods can be treated as a single aggregate if their prices move in parallel. To apply this result to the labor supply example, the real prices of all consumption goods must be perfectly collinear, which is inconsistent with price movements in the United States. Lewbel (1996) has described somewhat weaker conditions under which this form of commodity aggregation can occur, but it remains to be seen whether his result can be used to define a single aggregate consumption good.

In the absence of Hicks-Leontief aggregation, an alternative approach is to assume separability of the commodity for which there are data from all other goods. In the labor supply example, the structure of the indirect utility function (suppressing the demographic effects) must be of the form:

$$V_k = V(w_k, V_c(p, M_{kc}), F_k), \qquad (16)$$

where w_k is the wage, p is the vector of prices for consumption goods, M_{kc} is the total expenditure on consumption goods, and F_k is the level of full expenditure. In this model V_c is the indirect subutility function defined over goods and can be interpreted as the price index of the composite commodity. Preferences of this type can be recovered from a single labor supply equation because total spending on the two goods (leisure and consumption) must equal full expenditure.

This suggests that changes in welfare can be measured using a single demand equation as long as demand patterns are consistent with the separable structure of the utility function shown in (16). Is this a reasonable description of preferences? In general, preferences of this type require strong elasticity equality restrictions among the goods that comprise the composite commodity. Tests of weaker forms of separability usually reject it.[25] Note also that this form of separability does not actually solve the data problem that often motivates single equation methods. That is, (16) shows that the "income" variable to be included in the labor supply equation is full expenditure, but this requires information on the quantity of leisure consumed as well as the spending on goods and services.

Since single-equation methods require strong aggregation assumptions, multiple equation models are sometimes used. These are frequently based on the demands for goods and services, but the effects of leisure, public goods, and other nonmarket commodities are

[25] Recent examples include Diewert and Wales (1995) (in the production context) and Lewbel (1996). Barnett and Seungmook Choi (1989) point out that many tests of separability have low power and they illustrate the problem using Monte Carlo simulations.

typically ignored.[26] The welfare effects derived from these models measure changes in the subutility function defined over goods (V_c in equation (16)), rather than changes in the overall within-period utility level. If $M_c(p, V_c)$ is the expenditure function defined over goods consumption alone, a monetary measure of the change in the subutility level is:

$$EV_{kc} = M_c(p^0, V^1_{kc}) - M_c(p^0, V^0_{kc}), \quad (17)$$

where $V^1_{kc} = V_c(p^1, M^1_k)$ and $V^0_{kc} = V_c(p^0, M^0_k)$.

The accuracy of (17) as a welfare measure depends on the consistency of the assumption of separability. Evidence presented by Michael Abbott and Orley Ashenfelter (1976), Barnett (1979), and Martin Browning and Meghir (1991) does not support the assumption of separability of leisure from consumption. This suggests the importance of exploiting relatively new data sets that have information on both consumption and leisure in examining the welfare effects of policy changes.[27]

Even if the separability assumption is reasonable, what is the relationship between (17) and the change in within-period utility? Michael Hanemann and Edward Morey (1992) have shown that (17) will not accurately measure the overall change in utility, because changes in the prices and total expenditure will lead to reallocations between (in this particular application) leisure and consumption.[28] Jeffrey LaFrance (1993) has found the empirical magnitudes of such biases can be quite large, so that seemingly innocuous assumptions concerning commodity aggregation can lead to misleading conclusions concerning the impact of price changes on welfare.

Welfare of Whom? In practice, the data largely dictate the units of observation for which the welfare effects are calculated. If expenditure data are used, spending is almost universally aggregated over the individuals who make joint consumption decisions, and welfare is measured for the household. If inferences concerning well-being are based on labor supply information, the observational unit is taken to be the individual (and the influences of other members of the household are often ignored). Neither of these approaches are entirely satisfactory, and recent research has demonstrated the importance of analyzing the intrahousehold interactions of individuals.

Theoretical justification for modeling the household as the decision unit is attributed to both Paul Samuelson (1956) and Becker (1981). While both approaches lead to a representative agent treatment of the family, the assumptions concerning the interactions among members are quite different. Shelly Lundberg and Robert Pollak (1996) refer to Samuelson's framework as a consensus model in which the members of the household pool resources and make spending decisions that are consistent with an agreed upon social welfare function. Becker's model yields the same empirical implications, but he assumes that the household is controlled by an altruistic parent who allocates resources within the household so as to maximize utility subject to a common budget constraint.

Because individual-level consumption data are unavailable, the Samuelson-Becker representative agent assumption

[26] As examples, see Muellbauer (1974), Jorgenson, Lau and Stoker (1980, 1982), and, more recently, Frank Wolak (1996).

[27] This information is now available in the Consumer Expenditure Surveys in the U.S. and the Family Expenditure Surveys in the United Kingdom.

[28] A condition under which it is unbiased is that the consumer is constrained to consume the same amount of leisure before and after the policy change. In this case, the change in the utility defined over the consumption goods is the same as the overall change in utility under the rationed equilibrium.

is invoked, and welfare calculations are made for the household. Research over the last decade has suggested that this assumption may be inappropriate. Using data from low-income households, Lundberg (1988) found that the assumption of joint utility maximization on the part of households does not describe the labor supply behavior of husbands and wives. While she found little empirical support for the joint utility model, the specification that treated men and women as autonomous agents did not perform much better for the subsample of families with children.[29]

A growing body of evidence suggests that the joint utility model may be inappropriate for modeling the allocation of consumption across goods and services. An implication of the representative consumer model is that the source of income (such as that attributed to the husband or wife) should have no influence on expenditure patterns, since all income is pooled and spending decisions are made on the basis of a common utility function. Using Brazilian data, Duncan Thomas (1990) has found that the share of income contributed by the mother or father influences the nutrient intakes, survival probabilities and anthropometric outcomes of children.[30] Browning, et. al. (1994) find that the share of income contributed by women influences the share of expenditures on goods consumed exclusively by them.

If the joint utility model is inconsistent with spending patterns, how are individuals making these decisions and what are the implications for applied welfare economics? Marilyn Manser and Murray Brown (1980) and Marjorie McElroy and Mary Horney (1981) have characterized household decisions as the outcome of a bargaining process among members, although one model has little empirical support over another.[31] Extending the work of Pierre Chiappori (1988, 1992), Browning et. al. (1994) give conditions under which spending decisions can be modeled as the outcome of a household "sharing rule". This rule can be described as a two-stage allocation in which resources are initially divided between household public goods and the total spending of each member. In the second stage, each individual decides on the division of his or her own allocation across the various goods. On the basis of this sharing rule, Browning et. al. find that allocations among couples depend on their relative ages and incomes, as well as their level of wealth. Household demand behavior is, therefore, inconsistent with a model based on a representative consumer.

What are the prospects for measuring the welfare of individuals within the household? Given the lack of individual-level consumption data, the intrahousehold distribution must be inferred, and it is difficult to assess the consistency of the estimated distribution with the actual distribution. Even if the distribution were known, welfare measurement requires an assessment of individuals' valuations of the goods. For private goods, this can be done in the usual way. Many goods consumed by the household, however, are public so that welfare measurement requires knowledge of each individual's willingness to pay. This is obviously difficult to assess. It is also undoubtedly the case that members of a family or household have interdependent preferences, which adds an additional

[29] Paul Schultz (1990) and Bernard Fortin and Guy Lacroix (1997) also present evidence against the joint utility model of labor supply using, respectively, data from Thailand and Canada.

[30] He also reports (Thomas 1994) the effects of differences in the education levels of parents on the anthropometric outcomes of children for Ghana, Brazil, and the U.S.

[31] See also Lundberg and Pollak (1993).

complication to the calculations. Regardless, recent work has made important progress, and an analysis of the intrahousehold distribution of welfare is an important area for future research.

Data Requirements. Implementation of any of the methods used to measure welfare requires comprehensive micro level expenditure data. This is true regardless of the number of goods or the degree of separability that is assumed. If labor supply is modeled, full expenditure is needed but it is not reported in labor surveys such as the CPS. If only a single consumption good is considered (e.g. gasoline in Hausman and Newey 1995) and labor supply is exogenous, total expenditure must be reported. Often, income is used as a proxy for consumption, which generally results in biased estimates of the demand parameters and the associated welfare effects.

Estimation of the demand equations and the Hicksian surplus measures also requires price variation across the sample. As mentioned previously, this is a problem for most applications because cross sectional budget studies report the expenditures on commodities but not the prices paid for them. As a result, prices are obtained from other sources, and variation is introduced through regional or temporal movements. The *Consumer Expenditure Surveys* provide quarterly data from 1980 that can be linked with price series such as the components of the Consumer Price Index (CPI) or the implicit price deflators of personal consumption expenditures in the National Income and Product Accounts. Price variation can also be incorporated by linking micro demand equations with an exactly aggregated set of macro demand functions.[32] Heterogeneity in demand patterns (and the associated welfare effects) is identified using variation in the cross section data, while the price effects are estimated using time series variation in the aggregate model. Application of this approach, though, requires the assumption that every household faces the same prices. Regardless of the method, it is an understatement to suggest that the prices used in demand analysis typically embody measurement error.

Finally, to reiterate another point made previously, the nonparametric and semiparametric estimates of the welfare effects of policy changes are plagued by slow rates of convergence to the true underlying function. Precise estimates can be obtained only with large numbers of observations and substantial variation in the independent variables. Overparameterization is a similar concern with the flexible series expansions that have been proposed. This problem can only be circumvented with a sufficiently large number of observations on household demands that include information on the prices paid for the goods. I'm unaware of the existence of any such data which suggests that further application of these methods will require the parallel development of alternative data sources.

2.4 *Measuring Welfare Using Index Numbers*

The overlap between consumer's surplus (and the Hicksian variations) and index number theory has been noticed for some time (Hicks (1942)). The generalized equivalent variation (6) evaluates the change in utility as a monetary measure of the difference in utility while A. A. Konus' (1939) quantity index represents it as a ratio:

$$Q_k(p, V^0, V^1) = \frac{M(p, V^1, A_k)}{M(p, V^0, A_k)}\,.$$

[32] This approach was followed by Jorgenson, Lau, and Stoker (1982).

This can be estimated using the econometric methods described above, but index number proponents have typically followed a different empirical strategy.[33]

The econometric approach assumes a parametric form for the utility (or demand) function that is the same for all households. Observations on expenditure patterns are used to recover the utility function and measure changes in welfare that result from actual or simulated policies. Heterogeneity is accounted for by allowing preferences and demands to be functions of household characteristics while unobserved differences can be accommodated through the stochastic specification of the econometric model.

The index number approach makes weaker assumptions on preferences, but the data requirements are more stringent. The basic strategy is to avoid functional form assumptions on preferences and evaluate the relative levels of welfare using Samuelson's (1948) principle of revealed preference. Not only is this method nonparametric, but it also makes unnecessary the assumption that individuals have identical preferences. Let x_k^0 and p^0 be the quantity and price vectors in the reference period, and define x_k^1 and p^1 to be the corresponding vectors in the current period. If preferences are internally consistent, the individual is at least as well off in the base period if $p^{0\prime}x_k^0 \geq p^{0\prime}x_k^1$ and similarly in period 1 if $p^{1\prime}x_k^1 \geq p^{1\prime}x_k^0$. If neither condition holds, nothing can be said about the relative welfare levels in the two periods. This incompleteness is an obvious impediment to practical applications of this fully nonparametric approach to welfare economics.

Beyond a definitive statement as to the direction of the change in welfare, the magnitude is often of interest. For this purpose, an explicit form of the index needs to be specified. Rather than enumerate the advantages and disadvantages of the numerous possibilities, I illustrate the key issues using an index proposed by Malmquist (1953):

$$Q_{Mk}(x_k^1, x_k^0, u) = \frac{d(u, x_k^1)}{d(u, x_k^0)},$$

where $U(x)$ is the direct utility function and d is a distance function defined as:[34]

$$d(u, x) = \max_{\lambda} \{\lambda : U(x/\lambda) \geq u, \lambda > 0\} .$$

This index has the correct ordinal properties of being greater than unity if and only if utility is higher in period 1 but, because U is unknown, is inherently unobservable.

While a point estimate of Q_M is not possible without further assumptions on the form of the distance function (and, therefore, preferences), upper and lower bounds on the index can be calculated. These bounds represent index number analogs to the regions of indeterminacy associated with revealed preference methods. Diewert (1981) has shown that if the consumer is rational, the bounds on the Laspeyres version of the Malmquist index are:

$$\min_i x_{ik}^1/x_{ik}^0 \leq Q_{Mk}\,(x_k^1, x_k^0, u^0) \leq \frac{p^{0\prime}x_k^1}{p^{0\prime}x_k^0}\,,$$

where $u^0 = U(x_k^0)$. These bounds need not be tight. In fact, the lower bound can be less than one and the upper bound greater than one, implying that it is impossible to say whether welfare has increased or decreased.

With stronger assumptions on preferences, it is possible to make more definitive statements concerning the direction of the change in welfare using

[33] Comprehensive surveys of index numbers are provided by Diewert (1981, 1990) and Amartya Sen (1979).

[34] Deaton (1979) presents a detailed discussion of the distance function and its application to consumer theory.

index numbers. In a seminal paper, Diewert (1976a) has shown that a number of commonly-used indexes are "exact" in the sense of being ordinally equivalent to the underlying utility function for some representation of preferences. A Laspeyres quantity index, for example, is exact if preferences are of the fixed coefficient or Leontief form. Diewert defined indexes to be "superlative" if they are exact for a utility function that is second order flexible. Such indexes are appealing because the assumed form of the utility function can be interpreted as a local second-order approximation to any well-behaved function. While not fully nonparametric, this approach represents an attractive middle ground.

The Malmquist index is generally unobservable but under certain circumstances it can be calculated using only the observed prices and quantities in the base and reference periods. Assume the distance function is of the translog form and the reference indifference contour is a geometric average of the base and reference utility levels $u^* = (u^0 u^1)^{1/2}$. Diewert has shown that the following quantity index is exact (and also superlative) and can be calculated using only observations on prices and quantities in the two periods:[35]

$$\ln Q_{Mk}(x_k^1, x_k^0, u^*) = \frac{1}{2}\sum_{n=1}^{N}(w_{nk}^0 + w_{nk}^1)\ln\frac{x_{nk}^1}{x_{nk}^0},$$

where $w_{nk} = p_n x_{nk}/M_k$ is the budget share of the nth good for the kth household. For a particular assumption on preferences, therefore, it is possible to estimate the change in welfare without using econometrics. This procedure has greater data requirements, however, because the demands in both periods must be observed to calculate the index.

Given the correspondence between quantity indexes and the generalized equivalent variation, it is perhaps not surprising that Harberger's (1971) representation of consumer's surplus also has an index number analog. Using a Taylor's series approximation to an arbitrary utility function, he showed that consumer's surplus is approximately equal to:

$$\Delta CS_k = \sum_i p_i^0 \Delta x_{ik} + (1/2)\sum_i \Delta p_i \Delta x_{ik}, \quad (18)$$

where $\Delta p_i = p_i^1 - p_i^0$ and $\Delta x_{ik} = x_{ik}^1 - x_{ik}^0$. With repeated observations on household demands, this representation of consumer's surplus is easily calculated. The question is whether it is ordinally equivalent to the underlying utility as is, say, the Malmquist or Konus indexes.

Diewert (1976b) demonstrated that (18) identifies the direction of the change in utility as long as demands are in the region for which the weak axiom of revealed preference is conclusive. The problem arises in the region of indeterminacy where (18) can be made positive or negative by simply scaling the prices in period 1. Martin Weitzman (1988) attempted to resurrect Harberger's index by suitably normalizing the prices in period 1 but Diewert (1992) pointed out that his welfare measure requires the assumption of homothetic preferences. It is not surprising that this is essentially the same condition that was derived by Chipman and Moore (1976b).

The index numbers described to this point require only two sets of observations on households' demands. In the absence of assumptions on the form of the utility function, it is only possible to obtain bounds on the index that need not be particularly informative. S.N. Afriat (1967) and Hal Varian (1982) have

[35] See Diewert (1976a, pp. 123–24). This particular index number formula is attributed to Leo Tornqvist (1936) and can be interpreted as a discrete approximation to a Divisia index.

presented methods of welfare analysis that are fully nonparametric but are more likely to provide definitive conclusions concerning changes in welfare.

Using multiple observations on demands, it is possible to describe the set of commodity bundles that are strictly preferred to the base period demands x_k^0. This is the convex hull of the set of all points revealed preferred to x_k^0. The minimum cost of attaining a consumption bundle in this set (at specific prices) provides an upper bound to the expenditure function. Conversely, we can determine the set of all points that are revealed to be worse than x_k^0. The minimum cost of attaining a point in the complement of this set provides a lower bound to the expenditure function. Given the over- and under-estimates, it is possible to create interval estimates of the change in welfare in either difference or ratio form.[36]

The most obvious advantage of the index number approach is the fact that functional form assumptions on preferences are unnecessary. This precludes the possibility of erroneous inferences being drawn because of the misspecification of the form of the utility function. They are also fairly easy to compute (at least with two observations) and do not require the implementation of complicated optimization algorithms as are often required with econometric methods. The region of indeterminacy that results, however, is a consequence of these weaker assumptions. Even if the bounds on the indexes are tight, or one uses indexes that are exact for a specific functional form, other problems arise. Multiple observations on expenditures are required for each household, and these data are rarely available. In fact, no micro data of this type are available in the United States, so almost all applications use aggregate data in conjunction with a representative agent assumption.[37] Moreover, it is impossible to evaluate the effect of simulated policies since the demands under the alternative scenarios are not observed. This precludes the use of these methods to an important application of welfare economics. Note also that individual demands are assumed to be measured without error, which, given the condition of most budget studies, is an overly-strong assumption.[38]

2.5 *Extensions*

While welfare measurement has been examined in the context of changes in prices and total expenditure, these methods have general applicability and have been extended in several directions. In this section I take a brief look at three important applications to illustrate some of the additional empirical and conceptual issues.

Equivalence Scales. The model described to this point has focused on the effects of price and expenditure changes on the welfare of a given household. I now consider the impact of changes in household characteristics, which is an issue with several important public policy applications. To determine the appropriate benefit levels of transfers from the Food Stamp program or Aid to Families with Dependent Children (AFDC), we need to assess the extent to which "needs" change with household composition. The goal of such adjustments is to ensure that all poor women with children receive

[36] See, for example, Manser and Richard McDonald (1988), who used this approach to estimate bounds on cost-of-living indexes.

[37] Some researchers circumvent this problem by assuming that all individuals within narrowly defined demographic groups have identical preferences. See, for example, Melissa Famulari (1995).

[38] See, for example, the discussion of the Consumer Expenditure Surveys by Daniel Slesnick (1992) and how these survey correspond to other consumption data.

benefits sufficient to attain the same subsistence standard of living. Assessments of this type are also needed to make the poverty line comparable across households, to measure the costs of children and, in some countries, to adjust long term payments to unemployed workers and their families.[39] In each application, a fundamental welfare economic question is being asked: How much additional income is required for a household with certain demographic characteristics to attain the same welfare as a reference household?

The answer can be represented as either the difference or ratio of expenditure functions. If A_R is the vector of characteristics of the reference household, the additional expenditure required for a household with attributes A_k to attain utility V_R is:

$$\Delta W_k = M(p, V_R, A_k) - M(p, V_R, A_R), \quad (19)$$

which is analogous to the equivalent variation. In index form, household equivalence scales are defined by:

$$m_0(p, V_R, A_k) = \frac{M(p, V_R, A_k)}{M(p, V_R, A_R)}. \quad (20)$$

Because of the policy implications, a number of studies attempt to measure the welfare effects of changes in household characteristics such as family size. Many of these propose ad hoc adjustments that measure the "needs" of households in arbitrary ways. Some studies focus on the nutritional or food requirements of the family members, while others define the scale to be an arbitrary function of the number of persons.[40] For the purposes of welfare comparisons, such scales have no obvious interpretation since they are developed outside of a welfare theoretic framework. Rather than summarize the vast literature on equivalence scales (instead, see the surveys by Browning 1992 and Lewbel forthcoming), I restrict consideration to the empirical issues that arise in this particular application of welfare economics.

In measuring the welfare effects of changes in prices and expenditures, efforts have been directed towards incorporating price and expenditure effects flexibly in the demand functions. Demographic variables are often treated as an afterthought in an effort to account for heterogeneity. With equivalence scales, the issue is how to model demographic effects in a way that does not overly restrict preferences. Among the simplest methods of incorporating demographic variables is to deflate the demands and expenditure by a general equivalence scale, so that the Engel curve is:

$$x_k/m_0(A_k) = x(M_k/m_0(A_k)).$$

This method is attributed to Engel (1895) and is widely used because it is easily estimated using a single cross section (see the discussion by Deaton and Muellbauer (1986)). The demographic variables have the same effect on all commodities, which seems overly restrictive since the addition of a child is unlikely to have the same effect on alcohol consumption as it does on the demand for milk.

S.J. Prais and Houthakker (1955) generalized this approach in their seminal empirical study by allowing for different demographic effects across goods which yields Engel curves of the form:

[39] Lewbel and Richard Weckstein (forthcoming) point out that welfare comparisons of this type are also made frequently in the legal system in wrongful death cases.

[40] The equivalence scales implicit in the official poverty lines in the U.S. are based on the nutritional requirements of households and are described by Mollie Orshansky (1965). Brigitte Buhmann et. al. (1988) define equivalence scales to be a power function of family size, where the power coefficient indicates the economies of scale in consumption.

$$x_{ik}/m_i(A_k) = x_i(M_k/m_0(A_k)) \quad (i = 1, 2, ..., n),$$

where $m_i(A_k)$ is the commodity-specific equivalence scale for the ith good. This approach is plagued by identification problems because the n scales are being estimated with information on (n-1) freely varying commodity demands (with summability). Moreover, Muellbauer (1980) has shown that this method of incorporating demographic effects is consistent with rational consumers only if preferences are of the Leontief form and substitution effects are zero.

These problems can be resolved if one specifies the direct utility function (rather than the demands) to depend on per equivalent quantities. In this model, developed by A.P. Barten (1964), the indirect utility function depends on the prices scaled by the commodity-specific equivalence scales:

$$V(p, M_k, A_k) = V(p^*, M_k),$$

where $p^* = (p_1m_1(A_k), ..., p_nm_n(A_k))$ is the vector of "effective prices" for the kth household. This yields demands of the form:

$$x_{ik}/m_i(A_k) = x_i(p^*, M_k) \quad (i = 1, 2, ..., n). \quad (21)$$

This implies a direct correspondence between household compositional effects and price movements. A change in composition, such as the addition of a child, changes the relative (effective) prices of goods and services and induces substitution away from "child goods". There is also a scale effect due to the increased resources needed to maintain the standard of living. Note that it is possible for the substitution effect to outweigh the scale effect for some goods so that an additional child leads to a decrease in demand. This is widely regarded as an implausible feature of Barten's model and Muellbauer (1977) shows that data from the U.K. are inconsistent with this method of incorporating demographic variables in the demand functions. The equivalence between price and compositional movements is particularly useful, though, because it facilitates identification of the commodity-specific scales as long as the data provide sufficient price variation.

Pollak and Wales (1981) have distinguished demand models that introduce demographic effects by scaling prices (as in the Barten model) from those in which the demands (or budget share equations) are simply translated by a function of the demographic characteristics. In this latter case the translation parameter γ_i is interpreted as a fixed cost or subsistence level of consumption of the good:

$$x_{ik} = \gamma_i(A_k) + x_i(p, M_k - \sum_{j=1}^{n} p_j\gamma_j(A_k)) \quad (i = 1, 2, ..., n). \quad (22)$$

Gorman (1976) proposed preferences that incorporate both translation as well as scaling, and these preferences have the attractive feature that the excessive substitution found in scaling models is less likely.[41]

In many ways, the measurement of the welfare effects of changes in household characteristics appears exactly analogous to the evaluation of the impact of price and expenditure changes. In each case we are making an assessment of the change in utility as key variables change using observed demand patterns. While superficially similar, a number of aspects of equivalence scale estimation distinguish it from the equivalent or compensating variations.

Measuring the welfare effects of demographic changes introduces a normative element to the analysis that was previously absent. It is necessary to address questions such as how much

[41] Generalizations of these models have been proposed by Lewbel (1985).

more expenditure a family of five needs to be as well off as a married couple. This requires assumptions of interpersonal comparisons of well-being that are not empirically refutable. While arbitrary, interpersonal comparisons are unavoidable if one is to make any statement concerning the relative levels of well-being of different types of households.[42]

While economic theory provides guidance concerning the arguments of the demand function in measuring the welfare effects of price and expenditure changes, it is of little help in determining what equivalence scales should depend on. In most applications, attention has focused exclusively on the effect of family size on well-being. Since equivalence scales will generally depend on any set of characteristics (observed or unobserved) which influence demand patterns, omitting the influence of other demographic variables results in biased estimates of the parameters of the demand functions and the scales. In the United States, demand patterns differ significantly across households distinguished by gender, race, age of the members, and region of residence so that any assessment of needs based on family size alone is likely to be misleading.

Finally, Pollak and Wales (1979), elaborated further by Pollak (1991), contend that information on household expenditure patterns cannot be used to measure the welfare effects of changes in household characteristics. Demand models are estimated conditional on the existing demographic composition, while welfare comparisons require the evaluation of the change in the unconditional utility level. Consider a couple that has been trying to have children for years and is finally successful. With no change in total expenditure, the conditional preferences will indicate a decrease in utility because of the increase in consumption requirements. The measured welfare effect of this demographic change does not take into account how the couple feels about having children. If the unconditional utility function could be identified, it would indicate an increase in well-being.[43]

More formally, let $V(p, M_k, A_k)$ be the preferences conditional on the characteristics A_k. The unconditional preferences incorporate how the family feels about having these character istics and can be represented as $V^*(V(p, M_k, A_k), A_k)$. Both descriptions of preferences have different implications for equivalence scales but yield the same demand system. This observational equivalence suggests that, as Pollak and Wales argue, household equivalence scales cannot be identified from demand data without further information. Blundell and Lewbel (1991) demonstrate that with price variation the best one can do is estimate “relative” equivalence scales.

This problem has several possible resolutions. It has been suggested that the solution is to model the choice of those characteristics that are endogenous. If we assume that fertility outcomes always reflect choices made by individuals (a strong assumption), we might be able to measure how couples feel about having children by explicitly analyzing their rational choice of family size. While this identifies the unconditional welfare effect of additional children, it does not solve the identification problem for characteristics that are

[42] A detailed analysis of the degree of interpersonal comparability required for the development of equivalence scales has been considered by Blackorby and David Donaldson (1988b) and Lewbel (1989a).

[43] The Pollak and Wales critique has been examined in detail by Browning (1992) and Blundell and Lewbel (1991).

exogenous such as age, race or gender. For these characteristics, one might identify the welfare effect using psychometric data that asks individuals how they feel about being a particular age or race. While possible, it is not practical since such information is not readily available.

An alternative approach is to impose identification restrictions on the model. As an example, Blackorby and Donaldson (1988b) and Lewbel (1989a) have considered equivalence scales that are independent of the reference utility level. This condition holds if the expenditure function is of the form:

$$M(p, V_k, A_k) = B(p, A_k)C(p, V_k).$$

Blundell and Lewbel (1991) have shown that this restriction on preferences is sufficient to identify the equivalence scales using demand data alone, although they find little empirical support for it in data from the United Kingdom. Equivalence scales of this type have been estimated for the United States by Jorgenson and Slesnick (1987). Browning (1992) has shown that the intertemporal allocation of consumption can also be used to identify equivalence scales at least partially. These approaches, along with the development and integration of alternative data sources, are promising avenues for resolving the identification problem.

Intertemporal Welfare Measurement. The emphasis to this point has been on the measurement of welfare in a static, single-period framework. Even if intertemporal separability holds, a snapshot of the household's welfare could give a misleading picture since the effects of policies often last more than one period. The bias associated with static welfare analysis has been demonstrated empirically by James Poterba (1989,1991) in his examination of the incidence of indirect taxes. He showed that when the distribution of the tax burden is measured using current income as a proxy for household welfare, indirect taxes appear more regressive than when welfare is measured using household consumption (which approximates permanent income). This suggests that significant errors can be made by ignoring the "lifetime incidence" of policies.[44]

A more formal description of the biases associated with using the static equivalent variation to measure intertemporal welfare changes is presented by Blackorby, Donaldson and David Moloney (1984) and Michael Keen (1990). Intertemporal welfare effects are often represented as the discounted sum of the within-period equivalent variations. This will differ from the lifetime equivalent variation to the extent that households have the ability to substitute intertemporally. Let V_L be the maximum level of lifetime utility of a household that lives T periods when the profiles of prices and interest rates are $\{p_t\}$ and $\{r_t\}$ respectively. If the (optimal) time path of utility corresponding to V_L at these prices and interest rates is $\{V_{kt}\}$, the lifetime expenditure function can be represented as:

$$\Omega_L(\{p_t\}, \{r_t\}, V_L) = \sum_t g_t\, M(p_t, V_{kt}, A_{kt}),$$

where:

$$g_t = \prod_{s=0}^{t} (1 + r_s)^{-1},$$

and r_0 is equal to zero. As in the static framework, the lifetime expenditure function can be used to represent an exact measure of the change in lifetime welfare.

Let V_L^1 be the maximum level of life-

44 See also the evaluation of a value added tax considered by Erik Casperson and Gilbert Metcalf (1994).

time welfare when the profile of prices and interest rates are $\{p_t^1\}$ and $\{r_t^1\}$ and denote the corresponding time path of utility as $\{V_{kt}^1\}$. In the reference scenario, prices and interest rates are $\{p_t^0\}$ and $\{r_t^0\}$ resulting in a lifetime utility level of V_L^0 and within period utilities $\{V_{kt}^0\}$. An exact measure of the change in lifetime welfare, evaluated at the reference prices, is:

$$\Delta W_L = \Omega_L(\{p_t^0\}, \{r_t^0\}, V_L^1) - \Omega_L(\{p_t^0\}, \{r_t^0\}, V_L^0). \quad (23)$$

This is a straightforward extension of the static version of the generalized equivalent variation (6).

The change in lifetime welfare is not equal to the discounted sum of within period changes in utility. Equation (23) can be rewritten as:

$$\Delta W_L = \sum_t g_t^0[M(p_t^0, V_{kt}^1, A_{kt}) - M(p_t^0, V_{kt}^0, A_{kt})] + \sum_t g_t^0[M(p_t^0, V_{kt}^3, A_{kt}) - M(p_t^0, V_{kt}^1, A_{kt})], \quad (24)$$

where $\{V_{kt}^3\}$ is the optimal time path of utility required to attain lifetime welfare V_L^1 when prices and interest rates are $\{p_t^0\}$ and $\{r_t^0\}$. The first term in (24) is the discounted sum of the within period changes in welfare. The second term is the gain to the household that results from its ability to adjust the time path of consumption as the profile of prices and interest rates change. This is nonpositive because the rational household minimizes the cost of attaining a given level of lifetime welfare, implying that the discounted sum of the intraperiod changes in utility overstates the change in lifetime welfare. The magnitude of the bias is related to the ability of the household to adjust consumption intertemporally.

Since the present value of the within-period changes results in a distorted assessment of lifetime well-being, an alternative empirical model is required to measure the change in welfare accurately. Jorgenson, Slesnick and Peter Wilcoxen (1991) describe a model in which households are assumed to be infinitely-lived and linked to similar households in the future through preferences that are consistent with intergenerational altruism. The infinitely-lived household, referred to as a "dynasty", is assumed to maximize an additive intertemporal utility function of the form:

$$V_L = \sum_{t=0}^{\infty} \delta^t \ln V_{kt}, \quad (25)$$

where $\delta = 1/(1+\rho)$ and ρ is the subjective rate of time preference. The within period indirect utility function V_{kt} is represented by:

$$\ln V_{kt} = \alpha_p' \ln p_t + \frac{1}{2} \ln p_t' B_{pp} \ln p_t - D(p_t) \ln (M_{kt}/m_0(p_t, A_{kt})), \quad (26)$$

where $m_0(p_t, A_{kt})$ is the number of household equivalent members in period t. The unknown parameters of this utility function can be estimated in the usual way using information on household expenditure patterns.

Lifetime utility is maximized over the time path of utility $\{V_{kt}\}$ subject to the nonlinear budget constraint:

$$\sum_{t=0}^{\infty} g_t M_{kt}(p_t, V_{kt}, A_{kt}) = \Omega_k,$$

where Ω_k is the wealth of the household. This optimization implies a discrete time Euler equation of the type:

$$\ln V_{kt} = \frac{D_t}{D_{t-1}} \ln V_{kt-1} + D_t \ln \left(\frac{D_{t-1} g_t m_0(p_t, A_{kt}) P_t}{\delta D_t g_{t-1} m_0(p_{t-1}, A_{kt-1}) P_{t-1}} \right), \quad (27)$$

where D_t denotes $D(p_t)$ and P_t is a price index. The Euler equation can be used to obtain a closed form representation of the maximum level of lifetime utility and the intertemporal expenditure function. The change in well-being from a policy change is then easily calculated using (23).[45]

The evaluation of changes in intertemporal well-being introduces several additional empirical issues. As in the static model, the parameters of the within-period utility function can be recovered from demand equations that are estimated using household expenditure data. The intertemporal allocation introduces additional parameters related to the first stage allocation of total expenditure in each period. It is necessary to estimate the subjective rate of time preference as well as a parameter (or set of parameters) related to the intertemporal elasticity of substitution.

In the model described above, the Euler equation (27) can be used in conjunction with the within-period allocation of expenditures across goods to estimate the unknown parameters. This increases the data requirements because it requires a panel of households whose expenditures are recorded over time. Since such data are rare, in general, and nonexistent in the United States and the United Kingdom, repeated cross sections have been used to create synthetic cohorts to identify the intertemporal preference parameters. The time variation in the cross sections facilitates the estimation of the expenditure price and demographic parameters related to the within-period allocation of goods and services. These types of models are becoming more common in investigations of consumption behavior (see, for example, Blundell, Browning and Meghir 1994) among others) but have only rarely been used to measure welfare. A notable exception is the estimation of the intertemporal welfare effects of children by Banks, Blundell and Ian Preston (1994).[46]

Another issue that arises in these models is the appropriate treatment of demographic characteristics. As described in the previous subsection, the characteristics that affect demands are usually assumed to be exogenous and only their conditional effects are evaluated. In a life-cycle context, exogeneity is no longer a reasonable assumption and the fact that these households choose some of their characteristics needs to be accounted for in a framework that measures lifetime well-being. Note also that the preferences described above assume time separability. Further work is necessary to generalize these models to accommodate habit formation and durability and analyze the resulting implications for welfare analysis.

Labor Supply: Kinks, Corners and Notches. A common empirical application of welfare economics is the analysis of tax-induced changes in wages. Rather than measure the change in utility per se, the goal is to evaluate the deadweight loss or excess burden of the taxes.[47] Assume initially that budget constraints are linear and individuals face prices p, have a wage w_k, full expenditure F_k, and that the indirect utility function is $V(w_k, p, F_k, A_k)$. If p^0 and w_k^0 are the (fixed) pre-tax prices and wage, a measure of the efficiency loss

[45] The intertemporal equivalent variation has also been estimated by Don Fullerton and Diane Rogers (1993) in their examination of the lifetime incidence of taxes.

[46] See also the paper by Pashardes (1991) who uses somewhat different methods to measure intertemporal equivalence scales.

[47] Rosen (1978) was among the first to use micro data on labor supply to measure the deadweight loss due to labor taxation. A more recent example is Blundell, Meghir, Elisabeth Symons, and Ian Walker (1988).

due to distortionary taxation using the equivalent variation is:[48]

$$DWL_k = F(w_k^1, p^1, V^1, A_k) - F(w_k^0, p^0, V^1, A_k) - R(w_k^1, p^1, F_k, A_k),$$

where R is the tax revenue collected, $F(\,)$ is the full expenditure function, $h(\,)$ is the labor supply function, $p_i^1 = p_i^0 + t_i$ $(i = 1, 2, \ldots, n)$, $w_k^1 = (1 - t_0)w_k^0$, and V^0 and V^1 are the pre- and post-tax utility levels. This measure of the deadweight loss attributable to the taxation of goods and leisure (i.e. labor supply) is the amount, in excess of the tax revenue collected, the consumer would be willing to pay (at existing prices) to forgo all taxes. Put another way, it is additional revenue that can be collected, with no loss in utility, using lump sum rather than distortionary taxation:

$$V(w_k^0, p^0, (F_k - (DWL_k + R)), A_k) = V(w_k^1, p^1, F_k, A_k).$$

To measure the deadweight loss, it is necessary to recover the full expenditure function so the empirical issues are essentially the same as those addressed previously. However, several aspects of modeling labor supply distinguish it from the demand models that have been considered to this point. While the consumption of goods and services are aggregated over the members of the household, we frequently have disaggregated information on their individual labor supplies. This suggests the importance of modeling the interaction of members' labor supply decisions as in Hausman and Paul Ruud (1984), Michael Ransom (1987), Lundberg (1988) and Schultz (1990).

Labor supply also introduces the problem that only market labor is observed in most data sets. Household production issues are usually ignored, as are other uses of time (other than leisure) that could influence well-being.[49] This is particularly important in examining changes in welfare over time in the U.S., because of the increase in labor force participation of women in the 1970s. At first glance it may appear that well-being declined over this period because of the decrease in the consumption of leisure. This conclusion would obviously change if women were simply substituting market for nonmarket labor.

Models that incorporate labor supply are also distinguished by the fact that behavior appears discrete rather than continuous because the budget constraint facing most workers has a number of kinks and corners. Some workers, for example, choose not to participate in the labor market and their well-being must be evaluated at the corner corresponding to zero hours of work. Working has fixed costs such as child care, commuting costs and other occupational expenses, which not only introduce kinks to the budget constraint but nonconvexities as well. These costs induce individuals either to stay out of the labor market or to work enough hours to cover the associated expenses. The dependence of the tax rates on the level of earnings also has the effect of introducing additional nonlinearities.

There are two related empirical issues. First, one must initially decide on the appropriate econometric model to estimate the labor supply equations. The second problem is the recovery of preferences from these equations and the subsequent measurement of

[48] Alan Auerbach (1985) attributes this particular measure to Herbert Mohring (1971) and provides a general survey of the measurement of excess burden. See also the discussion by Peter Diamond and Daniel McFadden (1974) and J.A. Kay (1980).

[49] A notable exception is Patricia Apps and Ray Rees (1996) who employ time-of-use data and an explicit household production model to examine the within family distribution of welfare.

changes in welfare. Hausman (1985) (drawing on his previous work) addressed the first issue by developing methods that can be used to estimate unknown parameters when choices are made subject to nonlinear budget constraints. This requires the determination of the desired hours of work h^* for each worker including the kinks and corners. Hausman's algorithm exploits the feature that individuals are maximizing utility over budget constraints that are the union of convex sets.

Once the desired level of labor supply is determined, the next issue is to determine the wage and expenditure at which h^* should be evaluated. If the optimum occurs over the interior of a segment of the budget constraint, h^* is evaluated at the wage and expenditure level that describes that segment. If the optimum occurs at a corner or kink, the function must be evaluated at the virtual expenditure and prices that rationalize this quantity of labor supply. If the desired labor supply occurs at a kink corresponding to H_1 hours of labor supply and $l_1 = T - H_1$ hours of leisure, J.P. Neary and Kevin Roberts (1980) have shown that the corresponding virtual wage w_k^v and expenditure F_k^v are defined implicitly by:

$$H_1 = h^*(w_k^v, p, F_k^v, A_k), \qquad (28)$$

where $F_k^v = F_k + (w_k^v - w_k)l_1$. If the optimum occurs at an interior point of a line segment, rather than a corner, the virtual wage is the observed after-tax wage. Once the desired labor supply is represented as a function of the relevant wage and expenditure, the unknown parameters of the labor supply equations can be estimated in the usual way using minimum distance or maximum likelihood methods.

The utility level of the individual, whether at a corner or not, is evaluated at the virtual wage and expenditure:

$$V_k = V(w_k^v, p, F_k^v, A_k).$$

With the representation of utility as a function of the virtual wage and expenditure levels, a monetary measure of the change in well-being can be calculated in the usual way using the equivalent variation. Blundell, Meghir, Symons and Walker (1988) have used these methods to examine the welfare effects of changes to the United Kingdom tax code.

Although labor supply is the most common application, the methods used to deal with nonlinear budget constraints have general applicability. They can be used to analyze the welfare effects of any policies that are made in the presence of kinks or corners in the budget constraints. Bruno De Borger (1989), Manser (1987), Robert Schwab (1985) and Slesnick (1996) use similar econometric models to examine the welfare effect of in-kind transfers. The provision of these benefits produces a binding kink in the budget constraint for recipients who, with an equivalent cash transfer, would consume less of the good than is transferred in-kind. Similar issues arise in models in which demand behavior is discrete (such as with consumer durables or housing) or where a large number of individuals decide not to make a purchase.

3. *Measuring Social Welfare*

The conceptual issues underlying the measurement of household welfare are, for the most part, well understood. Consumption behavior reveals consumers' willingness to pay so that demand equations can be used to recover their expenditure functions and measure the equivalent or compensating variations. The focus of attention is not on *what* concept should be used to measure well-being but the empirical methods used to model demand behavior.

What emerges from this analysis are

estimates of welfare for households that are distinguished by their demographic characteristics. While these welfare estimates are of independent interest, the goal of empirical welfare analysis is often an assessment of the effects of policies on groups of households. We want to know whether tax reform hurts the poor more than the rich or if changes in transfer programs affect women and children adversely. Estimates of household welfare are an essential first step, but for welfare economics to be useful to practitioners, a method of aggregation must be developed.

Any effort to develop an index of group welfare must inevitably make normative judgements in which the gains to some are weighed against the losses to others. Should a policy be implemented if the welfare of the poor increases slightly but the welfare of the rich decreases dramatically? The answer ultimately depends on the extent to which well-being can be compared across the population and on the weights assigned to individual agents.

Due largely to economists' reticence to confront such normative questions, measures of group welfare used by practitioners have been developed in an ad hoc manner with little or no welfare economic justification. Rather than avoid the issue, a more appealing strategy is to confront it head on and make explicit the subjective judgements that are required to aggregate household welfare measures consistently into measures of social welfare. These assumptions take the form of choosing a representation for the household welfare function, stating the degree of measurability and comparability across households, and deciding the form of the social welfare function. A critical empirical issue is the extent to which social welfare judgements vary across these empirically unverifiable assumptions.

To illustrate the judgements that often remain implicit in policy analysis, consider per capita income as an indicator of social welfare. Income is used to represent individual welfare even though it bears little relation to any of the welfare functions described in Section 2. It is further assumed that income can be used to measure and compare well-being and that the social welfare function is a utilitarian index in which each member of the population is given equal weight. Although widely used, per capita income is no less arbitrary than an explicitly specified social welfare function. Over the remainder of this survey, I consider the issues related to social welfare measurement.

3.1 *The State of Practice*

Given a well-defined approach to the measurement of welfare at the micro level, how can the resulting indicators be aggregated to yield consistent rankings of social outcomes? There are four methods that are used to varying degrees.

Representative Agent Model. The easiest approach to finessing the aggregation problem is to assume that market demands are generated by a representative consumer. Under this assumption, the preferences of the consumer are revealed by aggregate demand patterns as long as the integrability conditions hold. The techniques described in Section 2 are applied to market data to obtain monetary measures of the change in utility for the representative agent.

This approach to social welfare analysis has several fundamental problems. Market demands need not be consistent with the assumption of a rational representative consumer. Hugo Sonnenschein (1972) has shown that even if every individual has demands that are consistent with utility maximization, aggregate demands need not satisfy any of the integrability conditions other than homogeneity of degree zero in prices and income and Walras' Law. Without further restrictions, there is no hope of using market demands to recover the preferences of the representative agent and measure social welfare.

What restrictions are necessary? Roughly speaking, the marginal income effects must be invariant across the population. Samuelson (1956) demonstrated that aggregate demands can be represented as those of a rational consumer if income is distributed so as to maximize a social welfare function. Given the disparities in the empirical distribution, this assumption can be rejected out of hand. Gorman (1953) described alternative functional form conditions that yield a representative agent. Consider a model in which aggregate demands are functions of prices and aggregate (or average per capita) expenditure:

$$\sum_{k=1}^{K} x_{ik}(p, M_k) = X_i(p, \Sigma M_k). \qquad (29)$$

Gorman showed that if individual consumers maximize utility, (29) holds if the micro-level demands are of the form:

$$x_{ik}(p, M_k) = \alpha_{ik}(p) + \beta_i(p)M_k, \ (k = 1, 2, \ldots, K).$$

As with Samuelson's condition, the marginal income effects must be constant across all consumers. If demands are zero when there is no income, preferences must be identical and homothetic. As mentioned previously, this restriction is inconsistent with longstanding empirical evidence.[50]

Aside from these empirical issues, additional conceptual problems are associated with using the representative agent model to measure social welfare. Distributional considerations are ignored from the outset, and these are often of paramount concern in examining the effects of different policies. Moreover, it's unclear what the utility function of the representative agent actually represents. Kirman (1992) presents an example in which the representative consumer prefers (aggregate) market basket A to B even though everybody in the economy prefers their allocation in B to that in A. This violation of the most basic principle of social choice suggests that the utility function of the representative consumer should not be used for policy analysis even in the unlikely event that market demands are consistent with the aggregation conditions.

Pareto Principle. In the early development of welfare economics, utility was considered to be both measurable and comparable across individuals. By the end of the 1930s, though, economists began to reject welfare judgements founded on interpersonal comparisons of utility. Lionel Robbins' (1932,1938) critique of this practice has been taken to be the watershed event that shifted the focus from utilitarianism to an exclusive reliance on the Pareto principle that requires only ordinally noncomparable utilities.[51] If everyone is better off under policy 1 relative to 2 then, under the Pareto principle, it is socially preferred as well. Such unanimity is the exception rather than the rule, and this method of welfare analysis has little practical usefulness.

Despite the fact that the ordering of social states is incomplete, the strict

[50] See Alan Kirman (1992) and Stoker (1993) for discussions of the evidence against the representative agent assumption. Less restrictive versions of the representative consumer model that incorporate additional features of the distribution of income have been presented by Muellbauer (1976) and Lewbel (1989b).

[51] Robert Cooter and Peter Rappoport (1984) provide an historical review of the debate of the appropriate conceptual basis for welfare economics.

application of the Pareto principle has remained a central element of welfare economics due, in part, to the parallel development of Kenneth Arrow's (1951) impossibility theorem. This famous result states that, under a set of "reasonable" conditions, the only social welfare function that ranks outcomes consistently is a dictatorship. This starkly negative finding became a major stumbling block to the empirical implementation of an explicit social welfare function. How is it possible to develop benevolent public policy if the only way to rank alternatives is to use the preferences of one person?

The Arrow impossibility theorem spawned a subfield of economic theory that focused on its robustness to modifications of the assumptions. Theorem after theorem demonstrated that modest changes to the axiomatic structure left the dictatorial result intact.[52] The inescapable conclusion was that if one precludes interpersonal comparisons of welfare, the only logically consistent foundation for welfare analysis is the Pareto principle. For economic theorists, welfare economics became completely summarized by the two Fundamental Theorems of Welfare Economics relating competitive equilibria to Pareto optimality.

Of course, this was bad news for practitioners. The Pareto principle can only rarely be applied in practice, and most would agree that a dictatorship is an inappropriate basis for the formation and evaluation of public policy. This left a gap between what theorists felt was a suitable framework for applied welfare economics and what policy analysts needed to provide a complete and consistent assessment of policies.[53]

Compensation Principles. As early as the 1930s, efforts were made to make the Pareto principle operational using compensation mechanisms. A typical criterion is for policy 1 to be judged an improvement over policy 2 if it is possible to reallocate goods in 1 to yield an allocation that is Pareto superior to 2. While this criterion has been used extensively, it introduces both conceptual and practical problems. It is possible to obtain what Blackorby and Donaldson (1990) refer to as "Scitovsky reversals" in which policy 1 is judged to be preferred to policy 2 and vice versa. Even if one eliminates from consideration allocations for which these reversals can occur, Gorman (1955) has shown that the ordering of policies under this criterion need not be transitive. These are serious shortcomings that limit the usefulness of this approach to ordering social states.

The Kaldor-Hicks-Samuelson approach described by Chipman and Moore (1971) provides a more stringent criterion for comparing outcomes. Policy 2 is preferred to policy 1 if, for any allocation of aggregate goods under policy 1, it is possible to find an allocation under policy 2 that is Pareto superior to it. Chipman and Moore have shown that this compensation principle does not suffer from the preference reversals and intransitivities that plague many compensation criteria. The stringency of the requirement for a welfare improvement, though, results in an ordering of social states that is often incomplete.

Ignoring the incompleteness, an important practical issue is how to make this specific compensation principle

[52] See Sen (1995) for a synthesis and review of the axiomatic basis for Arrow's famous theorem.

[53] It should be noted that while most of the profession rejected the feasibility of social welfare functions, Abram Bergson (1938) and Samuelson (1947) maintained the position that explicit social welfare functions were an appropriate basis for policy analysis.

operational. Clearly, it is not feasible to examine all possible lump-sum redistributions of goods across the population under each scenario. Efforts focused on aggregate index numbers as indicators of changes in "potential welfare" as revealed by the Kaldor-Hicks-Samuelson criterion. Let X^2 be the aggregate demands under policy 2 and define X^1 to be the corresponding demands under policy 1. Valuing the demands using the prices of policy 1, real aggregate expenditure has increased under 2 if:

$$p^{1\prime}X^2 \geq p^{1\prime}X^1. \tag{30}$$

The Paasche version of the increase in real expenditure is given by:

$$p^{2\prime}X^2 \geq p^{2\prime}X^1. \tag{31}$$

If either or both of these conditions reveal an increase in potential welfare, we have a relatively simple method of measuring changes in aggregate welfare.

Unfortunately, conditions (30) and (31) reveal increases in potential welfare (as indicated by the Kaldor-Hicks-Samuelson criterion) only under restrictive conditions. Chipman and Moore (1973, 1976a) have shown that (30) is only a necessary condition for an increase in potential welfare under policy 2. Both the Paasche and the Laspeyres indicators of the increase in real aggregate expenditure are sufficient conditions for an increase in potential welfare if and only if preferences are identical and homothetic.

Given these negative results, recent efforts have used summary statistics other than real aggregate expenditure to describe changes in potential welfare. Blackorby and Donaldson (1990) and Javier Ruiz-Castillo (1987) show that the sums of the individual equivalent or compensating variations reveal the direction of change in potential welfare under conditions similar to those derived by Chipman and Moore. This suggests that even if one accepts the Kaldor-Hicks-Samuelson criterion as the basis for measuring the change in welfare, there is no obvious method of implementing it without imposing very strong assumptions on preferences. Perhaps most important, Sen (1979) points out that the "New Welfare Economics" is, in a sense, irrelevant. If the compensation is not provided, or, even stronger, cannot be provided due to the unavailability of lump sum taxes and transfers, its impact on welfare is never realized and should have no influence on public policy. If it is provided, we can simply apply the Pareto principle.

Aggregate Surplus Measures. Rather than use the sum of the compensating or equivalent variations as indicators of potential welfare, an alternative approach is to define a function over the individual surplus measures as an explicit representation of the change in social welfare. Such an approach to aggregation was advocated by Harberger (1971), among others, in his effort to make consumer's surplus the standard tool for applied welfare analysis. At first glance, this appears to be a reasonable way to proceed. The equivalent variation provides an exact representation of the change in welfare for each individual, and would seem to be an ideal candidate for the arguments of a social welfare function.

At a conceptual level, such measures of social welfare are represented as natural extensions of the positive analysis of welfare measurement at the micro level. This is obviously not the case because aggregation necessitates normative judgements related to the treatment of unequally situated individuals. If a policy increases the welfare of a rich person by \$1000 and lowers the welfare of a poor person by \$500, has social welfare increased or decreased?

Simply summing the surplus measures, as is common, embodies a version of utilitarianism, ignores distributional concerns, and would indicate an increase in aggregate welfare. This and other approaches to aggregation requires assumptions concerning distributive ethics and there is no way around this issue. The differences in the various approaches lies in whether the assumptions are implicit or explicit.

Even if the ethical basis of the aggregation is explicit, there are reasons to avoid using surplus measures as arguments of social welfare functions. To see why, let $m(p,v) = (M_1(p,V_1), M_2(p,V_2), \ldots, M_K(p,V_K))$ be the vector of individual expenditure functions and define a social welfare function Ψ over the distribution of expenditure functions as:

$$E(p,v) = \Psi(m(p,v)),$$

where $v = (V_1, V_2, \cdots, V_K)$ is the vector of utility functions. This appears to be a reasonable approach to aggregation since the expenditure functions are exact representations of individuals' underlying preferences and Ψ can be specified to satisfy the fundamental principles of social choice.

The problem is that the expenditure function provides an exact representation of individual preferences for a fixed set of reference prices p. The choice of the prices used to "cardinalize" preferences is arbitrary, but the ordinal ranking of social states needs to be invariant to this choice. That is, the social welfare function should satisfy:

$$E(p,v) \geq E(p,v') \text{ if and only if } E(p',v) \geq E(p',v'),$$

for all possible v, v', p and p'. This is the property of "price independence" introduced by Roberts (1980a) who has shown that it will hold only under restrictive conditions. In the absence of restrictions of individual preferences, the social welfare function must be dictatorial. With no restrictions on the social welfare function, individual preferences must by identical and homothetic.[54] In addition, Blackorby and Donaldson (1988a) have shown that social welfare functions that use money metric utility functions as arguments need not be quasiconcave and, therefore, exhibit undesirable ethical properties.

3.2 *The Aggregation Framework: Social Choice Theory*

Since any method of aggregation necessarily involves subjective judgements, a reasonable way to proceed is to state explicitly the underlying ethical basis for the social ordering, and Arrow's (1951) axiomatic framework provides an ideal setting for doing this. Using the notation of Sen (1977) and Roberts (1980c), let R be a complete, reflexive and transitive ordering over the set of social states X. A social welfare functional f is a mapping from the set of individual welfare functions to the set of social orderings so that $f(u) = f(u')$ implies $R = R'$ where $u = (W_1, W_2, \ldots, W_K)$ is a vector of individual welfare functions. Using this notation, the axioms underlying Arrow's impossibility result are:

1. *Unrestricted Domain*(U): The social welfare functional f is defined for all possible vectors of the individual welfare functions u.

This seems like an innocuous assumption although it is worth noting that if the individual welfare functions are related to utility functions, the domain need not be unrestricted. In this case, the feasible set is limited to welfare functions that are consistent with the integrability conditions.

[54] These results have been extended by Blackorby and Donaldson (1985, 1990) and Slesnick (1991a).

2. *Independence of Irrelevant Alternatives* (IIA): For any subset A contained in the set of all social states X, if $u(x) = u'(x)$ for all x in A, then $R{:}A = R'{:}A$, where $R = f(u)$ and $R' = f(u')$ and $R{:}A$ denotes the social ordering over A.

Roughly speaking, the IIA assumption implies that the ranking of any two social states must be independent of a third. This axiom has been the subject of controversy and precludes social welfare functions that are based on, for example, rank-order voting. It also eliminates the possibility of incorporating principles of horizontal equity in which outcomes are evaluated relative to a reference social state such as the status quo.

3. *Weak Pareto Principle* (WP): For any x, y in X, if $W_i(x) > W_i(y)$ for all individuals i, then xPy where P indicates strict social preference.

This form of the Pareto principle implies that if everybody is better off in one state relative to another, it is socially preferred as well.

Perhaps the most important assumption underlying social choice is the assumption of the measurability and comparability of individual welfare functions. Motivated by the fact that demands are invariant to monotonic transformations of the utility functions that can differ across individuals, Arrow assumed that welfare functions are ordinally noncomparable. Any other assumption concerning the extent to which welfare can be measured or compared cannot be refuted by empirical evidence. To describe this condition formally, let L_k be the set of admissible welfare functions for the kth individual, and define L to be the Cartesian product of all such sets. Let Λ be the partition of L with elements that yield the same social ordering.

4. *Ordinal Noncomparability* (ONC): The set of welfare functions, Λ, that yield the same social ordering is defined by:

$$\Lambda = \{u' : W_k' = \varphi_k(W_k)\},$$

where φ_k is increasing and $f(u) = f(u')$ for all u in Λ .

Relative to alternative assumptions related to interpersonal comparability, ONC requires the least information on individual preferences, but imposes the most stringent invariance requirement on the social welfare functional.

If one accepts these seemingly reasonable conditions on social preferences, Arrow proved that the only possible social welfare functional is a dictatorship. This raises the question of the additional information that is required for a more inclusive approach policy analysis. The breakthrough occurred with a series of seminal books and papers by Sen (1970, 1973, 1977) who approached the problem of collective choice in a less nihilistic light. He viewed the Arrovian axioms as restrictive, rather than reasonable, because they rule out by assumption the ability of the analyst to weigh the gains of the winners against the losses of the losers. He demonstrated that relaxing the measurability and comparability assumptions expands the spectrum of possible social welfare functionals dramatically.

Sen's work changed the focus of research from demonstrations of impossibility to descriptions of the conditions under which a nondictatorial social welfare function is feasible. Others realized that practical application of welfare economics requires information based, either implicitly or explicitly, on the analyst's assessment of the relative impacts of policies on different individuals. The greater the (assumed) ability to measure and compare welfare levels, the wider the set of possible social welfare functions. Peter Hammond (1976) considered the case in which policy makers are able to rank welfare levels ordinally across individuals but cannot measure the magnitudes of the differences in well-being:

Ordinal Level Comparability (OLC). The set of welfare functions, Λ, that yield the same social ordering is defined by:

$$\Lambda = \{u' : W_k' = \varphi(W_k)\},$$

where φ is increasing and $f(u) = f(u')$ for all u in Λ .

Under OLC, welfare levels are unique up to monotonic transformations that are the same for every individual. Relative to ordinal noncomparability, this assumption imposes weaker invariance restrictions on the social ordering but requires more information since the analyst must now rank the relative standing of each member of the population. Maintaining variants of the other Arrovian assumptions, Roberts (1980b) showed that the set of possible social welfare functionals are "positional" dictatorships such as the case in which the worst-off individual is decisive or its lexicographic extension.[55]

If the analyst is unable to rank welfare levels but can determine that a change in policy hurts individual A twice as much as individual B, welfare levels are cardinal unit comparable and the social ordering must satisfy the condition:

Cardinal Unit Comparability. The set of welfare functions, Λ, that yield the same social ordering is defined by:

$$\Lambda = \{u' : W_k' = \alpha_k + \beta(W_k)\},$$

where $\beta > 0$ and $f(u) = f(u')$ for all u in Λ.

Under the assumption that changes in welfare are comparable, the social welfare functional is utilitarian and can be represented as the sum or weighted sum of welfare functions.[56] The assumption that changes in welfare are comparable has radically different implications for the ranking of outcomes compared to ordinal level comparability. While the utilitarian framework has no distributional concern, it is of paramount importance with the Rawlsian social welfare functional obtained under level comparability.

As the assumptions of comparability get stronger, the invariance restrictions on the social welfare functional get weaker, and the set of feasible social welfare functionals expands. Ordinal noncomparability assumes the least about the analyst's ability to measure and compare welfare levels but imposes the most stringent invariance requirements. If, at the opposite extreme, it is assumed that both levels and changes of welfare are comparable, the set of feasible social welfare functionals expands to those of the following canonical form:[57]

$$W(u) = \overline{W} + g(W_1 - \overline{W}, W_2 - \overline{W}, \ldots, W_K - \overline{W}),$$

where

$$\overline{W} = \frac{\sum_{k=1}^{K} W_k}{K},$$

and $g(.)$ is a linearly homogeneous function. Social outcomes are ordered on the basis of the sum of the mean and an index of deviations from the mean, so the social welfare functional incorporates concern for both efficiency and equity.

[55] That is, the worst-off, nonindifferent individual is decisive. See also Hammond (1976) and Sen (1977) for alternative characterizations of social welfare functionals related to the Rawlsian Difference Principle.

[56] For further details, see Roberts (1980c, p. 429). Eric Maskin (1978) provides an alternative characterization of utilitarianism.

[57] This is the assumption of cardinal full comparability and is similar to cardinal unit comparability except that the social ordering must be invariant to affine transformations that are the same for every individual. See Roberts (1980c) for further details.

The conclusion is that changing the measurability and comparability assumptions expands the set of possible social welfare functionals dramatically.[58] "Possibility" theorems were synthesized and extended in a series of papers by Roberts (1980b,c,d). By the early 1980s, theoretical results demonstrating the existence of logically consistent, nondictatorial social welfare functions had been established but remained unexploited empirically.

3.3 *Implementation Issues in the Measurement of Social Welfare*

Just as the theory of consumer behavior provides a framework for measuring individual welfare, social choice theory serves a similar function in measuring aggregate well-being. In evaluating alternative realized outcomes, the social welfare functionals suggest specific forms for the Bergson-Samuelson social welfare functions that can be used.[59] This requires that a number of additional practical issues be resolved before empirical implementation of the social welfare measures is possible.

As a way of illustrating the practical choices that need to be made in aggregating individual welfare functions, consider a widely used social welfare function that was originally proposed by Anthony Atkinson (1970):

$$W(u) = \left(\frac{1}{1-\rho}\right) \sum_{k=1}^{K} y_k^{1-\rho}, \qquad (32)$$

where y_k is the income of individual k, and the individual welfare function is taken to be $W_k = y_k^{1-\rho}/(1-\rho)$. In addition to the conditions of unlimited domain, independence of irrelevant alternatives and the weak Pareto principle, this utilitarian social welfare function requires the assumption of cardinal unit comparability.

The Units of Observation. Whose welfare functions should serve as arguments of the social welfare function? For a comprehensive assessment of aggregate well-being, (32) should be evaluated using all members of the population. The problem is that a large segment of the population (e.g. children and other dependents) has no sources of income but access to resources that enable them to maintain a reasonable standard of living. Using individual incomes or restricting the sample to those who receive income, as is typical in examinations of earnings distributions, undoubtedly distorts assessments of social welfare.

This problem could be resolved if we knew the resources available to each person or, better yet, their consumption. As a practical matter this type of information is almost never available. As a second best alternative, the unit of observation is typically taken to be the household and the distribution of resources among members is inferred. In the simplest version of this story, income or consumption is assumed to be divided equally among the members (a per capita model) or an optimal intrahousehold allocation is assumed. In the latter case, each member attains the same welfare and the household can be treated as a single autonomous agent. Both assumptions are obviously gross

[58] Of course, it is also possible to have a nondictatorial social welfare functional without interpersonal comparability if other axioms are relaxed. For example, if preferences are defined over a single dimension and are single peaked (violation of unlimited domain), the social welfare function is majority rule and aggregate preferences are those of the median voter. Sen (1977) and Roberts (1980c) describe in greater detail the possibility theorems that result when axioms other than those related to interpersonal comparability are relaxed.

[59] For further discussion of the link between the multiple profile framework of social choice described above and the single profile framework described by Bergson (1938) and Samuelson (1947), see Roberts (1980d).

simplifications and there is evidence (Lawrence Haddad and Ravi Kanbur 1990, Apps and Rees 1996, and others) that shows that such inferences can distort inequality and poverty assessments.

In light of the existence of household public goods and the absence of comprehensive data on within-household transactions, welfare measurement at the lowest level of aggregation (i.e. individuals rather than households) is far from straightforward. Until we learn more about within-household allocations of resources, the household is likely to remain the primary unit of observation in most empirical welfare studies. Note, however, that treating the household as the basic observational unit greatly complicates the problem of welfare comparisons because of the heterogeneity that arises from differences in composition. This issue is of central concern in the choice of the household welfare function.

Choosing a Household Welfare Function. In the absence of changes in the axioms of unlimited domain, independence of irrelevant alternatives or the weak Pareto principle, the social choice results described in the previous section suggest that some form of measurability or comparability is necessary to overcome Arrow's impossibility theorem. Where does this additional information come from? Can this requirement be met by choosing a particular welfare function? At first glance, it appears that the use of income as an argument of the social welfare function, as in (32), resolves the problem of measuring and comparing welfare levels. Income is cardinally measurable and is meaningfully compared across households. It is important to remember, though, that in this context income is to be interpreted as a measure of welfare. Without additional assumptions, it orders states uniquely up to monotonic transformations that may differ across households. Any representation of the welfare function, including income, money metric indexes or the indirect utility function, requires assumptions related to their measurability and comparability across the population. The statement that a household with twice the income of another is twice as well off is no less normative than a similar statement based on utility levels. The fact that income is measured in familiar monetary units doesn't diminish the difficulty of the exercise.

Given its necessity, is it possible to infer interpersonal comparability from household behavior? The answer is clearly no. We've known since the Ordinalist Revolution that demand functions enable us to identify utility functions that are unique up to monotonic transformations that can differ across households. For many economists, the failure to provide a "scientific basis" for interpersonal comparisons suggests that they shouldn't be invoked at all. As Sen (1995) points out, to do so limits the possibility of meaningful social choice and returns us to Arrow's impossibility result with its limited scope for policy analysis.

Social choice theory treats the household welfare function as an abstraction even though implementation of any social welfare function requires an explicit representation of W_k. Given the volume of work devoted to the measurement of Hicksian surpluses at the micro level, it is surprising that nominal income remains the most popular choice as a measure of well-being. In the notation introduced in the previous section, the welfare of the kth household is:

$$W_k = y_k. \tag{33}$$

This choice is implicit in the use of average household income as a measure of social welfare as well as in the majority of inequality and poverty studies that are based on the distribution of household income.

The assumption implicit in (33) is that a ten-person household is as well off as a couple with the same income. Given the disparity in the consumption requirements of these two households, this assumption may not be reasonable. To address this concern, a widely used adjustment is to deflate income by the number of persons in the household, say N_k, so that welfare is represented by:

$$W_k = \frac{y_k}{N_k}. \qquad (34)$$

Consumption requirements are assumed to increase linearly with family size, which may be true for some goods but certainly is not true for household-level public goods such as housing. To account for economies of scale in consumption, Buhmann et. al. (1988) suggested a slightly different welfare function:[60]

$$W_k = \frac{y_k}{(N_k)^{\eta}}, \qquad (35)$$

where the parameter η determines the scale economies. If it is equal to one, then the household has no economies of scale (as in the per capita model) while a value of zero implies complete economies (as in the household model). This income-based welfare function has the advantage of simplicity but is ad hoc and has no apparent welfare-theoretic foundation. Note also that in accounting for heterogeneity, the effects of other family characteristics (such as the age and gender of the family members) are ignored.[61]

The widespread use of income as a welfare measure is related to several practical considerations. Income data are readily available in most countries and, at least in the United States, are provided at annual frequencies. Its use also obviates the need to estimate a model and recover the underlying utility or expenditure function. These advantages come at great cost, however, because income-based methods produce a number of distortions in both the levels and distributions of well-being.

One source of bias arises from the fact that judgements are being made in the absence of information on prices. If welfare is related to the utility function, a social ordering defined over income will coincide with another that is based on the indirect utility function (and, therefore, is a function of prices) only if household preferences or the form of the social welfare function are severely restricted.[62] Moreover, in a static context, the appropriate welfare indicator should be a function of total expenditure rather than income. The permanent income and life cycle hypotheses suggest that (annual) income is likely to be an inaccurate proxy for total expenditure because of intertemporal consumption smoothing. Households with temporarily low incomes will maintain their standard of living through borrowing or dissaving and exhibit high consumption to income ratios. The reverse will hold for households with unusually high incomes. For both types of households, income will typically be an inaccurate reflection of their standard of living.

Many of these problems can be resolved if welfare is represented by a money-metric utility function such as:

$$W_k = \frac{M_k}{P_k(p^0, p, V_k, A_k)m_0(p^0, V_k, A_k)} \qquad (36)$$
$$= M(p^0, V_k, A_R),$$

[60] See also Fiona Coulter, Frank Cowell and Stephen Jenkins (1992).

[61] The importance of accounting for characteristics other than family size has been demonstrated by Banks and Paul Johnson (1994).

[62] See Roberts (1980a) and Alan Slivinski (1983).

where $V_k = V(p, M_k, A_k)$ is the indirect utility function and A_R is the vector of attributes of a reference household. These monetary measures of welfare were first used as arguments of a social welfare function by Muellbauer (1974) who estimated the expenditure function econometrically using a linear expenditure demand system. Subsequent applications of this general approach have been used in a variety of contexts by Apps and Elizabeth Savage (1989), Deaton and Muellbauer (1980a), King (1983), and Martin Ravallion and S. Subramanian (1996).

The price index P_k in (36) accounts for the fact that price changes have different effects on different households depending on their expenditure patterns and is defined by:

$$P_k(p^0, p, V_k, A_k) = \frac{M(p, V_k, A_k)}{M(p^0, V_k, A_k)}.$$

If the prices of necessities increase relative to luxuries, households that consume large proportions of these goods (e.g. the poor) will experience larger increases in the cost of living and, all other things equal, will attain lower welfare levels.

In contrast with the ad hoc formulations used in (33)–(35), equivalence scales in (36) are designed to address the welfare theoretic question of the additional expenditure required to maintain a given level of well-being as household characteristics change. Recall from (20) that household equivalence scales are defined as the expenditure, relative to a reference household, necessary to attain utility level V_R:

$$m_0(p, V_R, A_k) = \frac{M(p, V_R, A_k)}{M(p, V_R, A_R)}.$$

If the equivalence scale is evaluated at the actual utility level V_k, per equivalent expenditure provides a reasonable basis for comparing the welfare of households with different compositions. If a family of four needs three times the expenditure of a single individual to attain the same utility, then (ignoring the differential effect of prices) we know that it is worse off with \$40,000 compared to the individual's \$20,000 and, moreover, would need another \$20,000 to be as well off.

Rather than account for heterogeneity through household size alone, equivalence scales depend on additional characteristics that influence demand patterns such as the age, race and gender of the members. It is worth repeating that using equivalence scales (derived from household expenditure patterns) for welfare comparisons remains an issue of some controversy dating back to Pollak and Wales (1979). What is lacking from the debate is a discussion of how one might compare well-being and develop aggregate measures of welfare without equivalence scales. Simply stating impossibility without an alternative proposal precludes the application of welfare economics to key elements of policy analysis including poverty and inequality studies, the analysis of growth, and the assessment of the impacts of taxes and transfers.[63]

The money-metric utility function (36) provides an exact representation of households' preferences. As long as the expenditure function is estimated using observed demands, it is consumption-based, incorporates price effects as well as preference heterogeneity. As an argument of a social welfare function, however, it suffers from the same problems as the aggregate surplus measures identified by Roberts (1980a) and Blackorby and Donaldson (1988a). Money-metric utility functions require the (arbitrary) choice of a reference

[63] Notable exceptions are the papers by Atkinson (1992) and Atkinson andFrancois Bourguignon (1987).

price vector and the ranking of social states need not be invariant to this choice. Social welfare functions defined over their distribution need not be quasiconcave which yields an ordering of outcomes with adverse implications for distributive ethics.

An alternative approach proposed by Jorgenson and Slesnick (1984) is to use an indirect utility function to represent household welfare:[64]

$$W_k = \ln V(p, M_k, A_k) = \alpha_p{}'\ln p + 1/2\ln p' B_{pp} \ln p - D(p)\ln(M_k/m_0(p, A_k)), \quad (37)$$

where α_p and B_{pp} are unknown parameters that are estimated econometrically. As with Muellbauer's model, this measure of well-being is a function of prices, the level of total expenditure, and the demographic characteristics of the household. Household welfare is not converted into monetary units using the expenditure function so the associated problems are avoided. Note, though, that the measurability issue does not go away, and it is still necessary to make assumptions concerning the cardinality and comparability of this indicator across households.

To this point, I have restricted the focus to social orderings that depend only on welfare functions that are themselves functions of the material well-being of households. Sen (1977) has argued that such a framework is unduly restrictive because it precludes the use of "non-welfare" information in ranking social states. He believes that social welfare should depend not only on the utility derived from consumption but also on characteristics that are not arguments of the welfare function. Nonwelfarist orderings of outcomes could be based on societal attributes such as the degree of personal liberty, the level of social justice and the ability of individuals to live in the absence of disease and malnutrition.

To accommodate this broader definition, Sen (1985, 1987) developed a theoretical framework in which the utility derived from consumption ("welfarism") is replaced by a function of individuals' "capabilities" such as an individual's life expectancy, health status, personal liberty and the like. Commodities are consumed for the characteristics they provide which, in turn, are inputs to the production of capabilities or "functionings" from which well-being is derived. Food is consumed to provide the characteristic related to nutrition which gives the capability of living in the absence of malnutrition. The medical care characteristic provides the capability of living a long life and the transportation characteristic provides the capability of physical mobility. In this context, consumption is best viewed as a means to an end rather than an end itself.[65]

It is difficult to imagine how this approach might be implemented empirically to provide a comprehensive welfare measure. Individuals' capabilities are not always the result of revealed preference so we have little prospect for measuring individuals' valuations of these capabilities. For example, it might be important to know how they trade off material well-being for, say, life expectancy. This can only be answered in an arbitrary way since data provide little information on this issue. Most empirical applications examine each capability separately but make no attempt to aggregate the joint effects into an overall measure of welfare. Nevertheless, it is difficult to deny the importance of

[64] See also Deaton (1977) and, more recently, Banks, Blundell, and Lewbel (1996) and David Newbery (1995).

[65] For further elaboration of this point, see Sudhir Anand and Ravallion (1993).

these issues in the assessment of social welfare.

Incorporating Principles of Equity. The axiomatic social choice framework often yields a canonical representation for the social welfare function. Once the household welfare function is chosen, the specification is completed by choosing an explicit functional form for the social welfare function. As with the issue of the measurability and comparability of welfare, the final decision on the functional form has important ethical implications related to the principle of vertical equity.

The concept of vertical equity is often identified with the notion that, all other things equal, more egalitarian distributions are preferred to those that are more dispersed. This condition is typically imposed by requiring the social welfare function to be quasiconcave, which implies that a convex combination of two distributions yields a level of social welfare that is not lower than that attained at the worst of the two. The condition of quasiconcavity does not narrow the set of feasible social welfare functions sufficiently, and includes functions as ethically diverse as the utilitarian specification and the Rawlsian maximin criterion. There is still flexibility on the part of the analyst to incorporate different levels of concern for equity by choosing social welfare functions with different degrees of curvature.

The weight one places on equity can be adjusted systematically using a social welfare function of the form:

$$W(u) = \left(\frac{1}{1-\rho}\right)\sum_{k=1}^{K} W_k^{1-\rho}. \qquad (38)$$

This representation coincides with that proposed by Atkinson (1970) when household welfare is represented by the level of income. In addition to the conditions of unlimited domain, independence of irrelevant alternatives and the weak Pareto principle, the social welfare function in (38) requires three additional restrictions. The social ordering must be invariant to the scaling of household welfare functions, and it must satisfy the conditions of anonymity and separability.[66]

The sensitivity of the social welfare function to disparities in well-being is related to the parameter ρ. A positive value ensures concavity and higher values imply greater concern for equity. In the limit, a value of ρ equal to zero yields a utilitarian social welfare function and the Rawlsian maximin welfare function is obtained as ρ approaches infinity. The choice of this parameter is arbitrary and reflects the analyst's aversion to inequality in ranking outcomes. In most applications, the social welfare function is evaluated at a number of different values to assess the sensitivity of policy conclusions to different levels of concern for vertical equity.

An alternative method of characterizing concern for vertical equity is to describe the effect of frictionless transfers of income or expenditure on the measured level of social welfare. Hammond (1977) described a distribution $\{M_k\}$ to be "more equitable" than $\{M_k'\}$ if:

(*i*) $M_i + M_j = M_i' + M_j'$

(*ii*) $M_k = M_k' \qquad (\text{for} \quad k \neq i, j)$

(*iii*) $V(p, M_i', A_i) > V(p, M_i, A_i) > V(p, M_j, A_j) > V(p, M_j', A_j).$

[66] This measurability and comparability assumption is referred to as "cardinal ratio scale comparability". Anonymity requires that permutations of assignments of welfare to households leaves the ordering unaffected. As in most contexts, the assumption of separability is quite strong. The normative implications in social choice theory is that individuals who are indifferent between outcomes have no influence on the social ordering. See Roberts (1980c) for further elaboration of the axiomatic basis of this social welfare function.

In other words, a distribution is more equitable than another if it results from a progressive transfer of expenditure from one household to another. He defined a social welfare function to be "equity regarding" if it is always larger for a more equitable distribution.

Given differences in the composition of households, it might be desirable to assign different weights to households with different characteristics. We may, for example, want to give a couple twice the influence of a single individual in ordering social states. The ultimate choice, which reflects a personal value judgement on the part of the analyst, influences the transfer sensitivity of the social welfare function in a significant way. The requirement that the social welfare function be equity-regarding can help determine the final choice of the weights and, simultaneously, make explicit the normative implications of this choice.

To illustrate, assume that the social choice axioms imply a general utilitarian social welfare function:

$$W(u) = \sum_{k=1}^{K} a_k W_k.$$

If the welfare function is taken to be of the type (37) and the weights are normalized to sum to one, then the requirement that the social welfare function be equity-regarding implies weights of the form:

$$a_k = \frac{m_0(p, A_k)}{\Sigma m_0(p, A_k)}.$$

For these representations of individual and social welfare, the condition that the social welfare function be equity-regarding implies identical weights for households with the same demographic characteristics. Moreover, these weights correspond to the number of adult equivalent members in each household rather than the number of persons.

Data Issues. Since household welfare functions are critical inputs to social welfare measurement, one needs all of the data necessary to estimate these functions. This includes information on spending patterns, the prices paid for goods and services and the demographic characteristics that influence consumption patterns. The aggregation of the household welfare functions adds several additional data requirements.

The most important requirement is that of comprehensive coverage over the set of households for which welfare is to be measured. While this seems obvious, it is often difficult to achieve because many surveys restrict the sample coverage. In the early 1980s, for example, the *Consumer Expenditure Surveys* limited the sample to urban households, which precludes its use for an assessment of the standard of living and its distribution across the entire population. Surveys in developing countries frequently do not include the rural sector, which is a particularly serious problem because this sector often represents a large segment of the population.

The absence of price information in the surveys creates special problems for the measurement of social welfare, inequality and poverty. Regional price variation in the U.S. is important for many goods (particularly housing), and ignoring its effects could influence social welfare judgements. Most empirical work links micro data with national price series on different types of goods so cross sectional price variation is ignored. Access to more disaggregated information on prices will enhance our ability to measure social welfare, although it remains to be seen whether fundamental conclusions concerning distributional issues will be affected.

Since welfare comparisons are critical to the calculation of social welfare, future efforts should focus on the development of data sources that facilitate the identification of the (unconditional) effect of demographic variables on household welfare. These data could include information on the intertemporal allocation of consumption, where progress has been made using synthetic cohorts to identify preference parameters. It is not clear whether this approach is helpful in the current context, since aggregation across age cohorts is required, and results in a loss of information that would identify the welfare effects across households distinguished by other characteristics. Subjective assessments of well-being and further information on the intrahousehold allocation of resources are also likely to provide important information in measuring household welfare.

3.4 *Applications of Social Welfare Analysis*

The aggregation issue in welfare economics lies at the core of some of the most important problems in applied economics. I present three examples as a means of illustrating the practical problems associated with the various applications of social welfare analysis.

Policy Evaluation. Following Harberger's (1971) lead, the sum of the equivalent or compensating variations is commonly used to represent the change in social welfare. While this has the appearance of being a positive measure of the change in aggregate welfare, it is no less normative than methods based on explicit social welfare functions. The sum of Hicksian variations depends on the distribution of well-being and the underlying ethical assumptions are often ambiguous. As pointed out several times already, problems are also associated with using individual expenditure functions as arguments of a social welfare function.

Once the social welfare function is specified, policy evaluation is a relatively straightforward exercise. A potential complication arises from the fact that social welfare need not be measured in monetary units and could be difficult to interpret. Money measures occur naturally if the arguments of the social welfare function are income or some other money metric but will not occur if utility functions such as (37) are used. In this case, social welfare can be converted into equivalent monetary units using Pollak's (1981) social expenditure function:

$$M(p, W) = \min \{M\colon W(u) = W, \Sigma M_k = M\}.$$

This function is exactly analogous to its micro level counterpart and is the minimum level of aggregate expenditure required to attain a given social welfare contour. The social expenditure function provides an exact monetary measure of social welfare, so the change in welfare that results from moving from policy 0 to policy 1 is:[67]

$$\Delta W = M(p, W^1) - M(p, W^0).$$

This general approach has been used in a variety of contexts. Deaton (1977) recovered the indirect utility function underlying an estimated linear expenditure system and examined the social welfare impact of indirect taxes in the United Kingdom. King (1983) defined a social welfare function over the distribution of indirect money-metric utility functions and measured the effect of changing the tax treatment of owner-occupied housing in the U.K.[68] Jorgenson and Slesnick (1985a,b) developed a

[67] Note that while the monetary measure is ordinally equivalent to the underlying social welfare function, the precise value of the money metric depends on the choice of reference prices .

[68] King recognized the problems with using money metric utility functions as arguments of the social welfare function and appropriately restricted consideration to a model with homothetic preferences.

model for analyzing the effects of energy taxes using a social welfare function that was defined over translog indirect utility functions and converted to monetary equivalents using the social expenditure function.

In these latter papers, we developed a decomposition of the change in social welfare into efficiency and equity components. To describe it, define $W_{\max}$ to be the potential level of welfare that is attained when aggregate expenditure is reallocated so as to eliminate disparities between households. If the social welfare function is equity-regarding, social welfare is maximized at this distribution for a fixed level of aggregate expenditure. The monetary equivalent of the maximum level of social welfare is $M(p, W_{\max})$ and the change in efficiency is:

$$\Delta EFF = M(p, W^1_{\max}) - M(p, W^0_{\max}).$$

Differences between the maximum level of social welfare and that attained at the existing distribution, $[M(p, W) - M(p, W_{\max})]$, is a monetary measure of the loss in social welfare due to an inequitable distribution of household welfare. A measure of the change in equity can be represented as:

$$\Delta EQ = [M(p, W^1) - M(p, W^1_{\max})] - [M(p, W^0) - M(p, W^0_{\max})],$$

and the change in social welfare is the sum of the change in efficiency and the change in equity:

$$\Delta W = \Delta EFF + \Delta EQ.$$

Implementation of this approach requires observations on the effects of policies on all segments of the population. Except in the unusual case in which we have a true natural experiment, it is not possible to discern the isolated (i.e. ceteris paribus) effects of policies using the data sets that are typically available. As a result, the effects are simulated and the behavioral responses (and associated welfare effects) are estimated using an econometric model. For this type of analysis it is not generally possible to use a non-econometric index number approach since, in this framework, the responses to changes in policy cannot be predicted.

The method of simulating these policies, therefore, plays an important role in the calculations. For ease of implementation this is most often done in a partial equilibrium framework where the broader consequences of a policy change are ignored. To examine the effects of changes in indirect taxes, for example, complete pass-through is often assumed and the effect is to increase the fixed producer prices by the amount of the tax. Over the last two decades, computable general equilibrium models have grown in popularity as a means of simulating alternative policies.[69] While these models are useful in assessing how prices and levels of income and consumption change, most are not sufficiently detailed to model endogenous changes in the distribution. These are undoubtedly important in measuring the social welfare effects of many of the policies of interest.[70]

After all is said and done, our confidence in policy recommendations ultimately depends on the robustness of our conclusions. Since equity concerns are based on subjective judgements, the effects are often evaluated for social welfare functions that exhibit different

[69] Early examples of computable general equilibrium models used for this purpose were provided by John Shoven and John Whalley (1972) and Edward Hudson and Dale Jorgenson (1974). More recent applications include Auerbach and Laurence Kotlikoff (1987), Jorgenson and Kun-Young Yun (1990) and Fullerton and Rogers (1993) among many others.

[70] Fullerton and Rogers (1993) provide the only computable general equilibrium model that I am aware of that models endogenous changes to the personal distribution of income.

sensitivity to changes in the distribution. This is typically done by evaluating social welfare for different values of a curvature parameter such as ρ in (38). Less common is an assessment of the role of the econometric model in the final conclusions. The level of social welfare is determined using statistical estimates of the household welfare function which, of course, have sampling distributions. Given a sufficiently large sample of households, the "delta method" could be used to estimate asymptotic standard errors for the change in social welfare. What is often overlooked, though, is that the general equilibrium simulations are themselves based on econometric estimates and the projections of outcomes of policies will likewise have sampling distributions which will affect the precision of the results.

A number of recent efforts have focused on reforms that are only small perturbations of the policy instruments. The justification for this narrower focus is that reforms are restricted (due to, for example, political considerations) so that large changes in, say, tax variables are infeasible. This has become known as the "tax reform" literature and has been used to describe directions of tax changes that improve social welfare. Ehtishan Ahmad and Nicholas Stern (1984) considered the reform of indirect taxes in India and measured the change in social welfare that results from tax-induced changes in prices.[71] More precisely, let social welfare be defined by $W = W(W_1, W_2, \ldots, W_K)$ where $W_k = V(p, M_k, A_k)$. Using Roy's Identity, the gradient of the change in social welfare resulting from price changes is:

$$\frac{\partial W}{\partial p} = -\sum_{k=1}^{K} \lambda_k x_k \,, \tag{39}$$

where λ_k is the social marginal utility of income for the kth household:

$$\lambda_k = \frac{\partial W}{\partial V_k} \frac{\partial V_k}{\partial M_k}.$$

It appears that for small changes in prices, the welfare effects can be approximated by (39) without estimating demand functions econometrically. Only observed quantities are required along with a specification of the relevant distributive ethics embodied in each household's social marginal utility of income. In fact, this initial impression is illusory. The social marginal utility of income will generally be a function of preference parameters which must be estimated. As Banks, Blundell and Lewbel (1996) point out, the assumption that the λ_k depend only on income imposes the condition of price independence, which is empirically untenable. They show, in addition, that the first order approximations provided by (39) can be quite inaccurate for the price changes that are typically of interest and that inclusion of higher order terms requires demand system estimation.

The Social Cost of Living. Estimates of the cost of living are among the most important statistics produced by the federal government. They are used to adjust wages, Social Security benefits, pensions and government transfers in an effort to preserve recipients' standard of living. Tax brackets are indexed to avoid what is commonly known as "bracket creep". Relatively minor biases in the price index used for these adjustments can have enormous impacts on fiscal policy. The key question is which index should be used for this purpose?

Over the years, the CPI has been used extensively and virtually exclusively as the government's estimate of the cost of living, despite increasing

[71] See also the applications by Newbery (1995) and Joram Mayshar and Shlomo Yitzhaki (1995).

evidence of systematic biases.[72] A panel of experts convened to study the CPI concluded that the inflation bias is approximately 1.1 percent per year.[73] This is in addition to permanent errors that were previously introduced in the estimate of the price level that can be attributed to the treatment of owner-occupied housing. One source of error arises from the fact that the CPI is a Laspeyres index that uses fixed weights in averaging the inflation rates of the individual goods, so substitution is not taken into account as relative prices change. Attention has also focused on the biases that result from the treatment of quality changes, the introduction of new goods, the treatment of elementary goods and statistical problems related to the sample rotation.[74]

These problems are of a mechanical nature and are largely unrelated to more serious conceptual issues related to the measurement of the cost of living. In most applications, the central issue is the amount nominal wages, benefits, or expenditures need to change to maintain a constant level of well-being. This is a basic welfare economic question that the CPI is ill-equipped to answer. In order for the CPI to be interpreted as a cost-of-living index, one must make the assumption that aggregate expenditure patterns are consistent with a representative consumer whose preferences are of the Leontief fixed coefficient form.[75]

A sound conceptual framework for cost-of-living measurement was provided by Konus (1939) who defined a price index at the micro level to be the relative cost of attaining a utility level V as prices change:

$$P_k(p^1, p^0, V, A_k) = \frac{M(p^1, V, A_k)}{M(p^0, V, A_k)},$$

where p^1 is the vector of prices in the current period, and p^0 is the vector of prices in the base period. By construction, this index answers the question of the ratio by which expenditure must change to maintain a household's well-being.[76]

While micro level price indexes may be of independent interest, of primary concern for policy analysis is the cost of living for a group of households. The most common approach, and the one implicit in the use of national price indexes, is to assume that market demands are generated by a representative consumer and the social cost of living is estimated using the same methods that were designed for households. I have already described the many problems associated with the application of the representative agent model to welfare economic problems. What, then, is the best way to measure the cost of living for groups?[77]

Prais (1959) proposed a "plutocratic" index as a weighted sum of household-specific price indexes where the weights are households' shares of aggregate expenditure:

$$P^P(p^1, p^0, u) = \frac{\Sigma M(p^1, V_k, A_k)}{\Sigma M(p^0, V_k, A_k)} = \Sigma s_k P_k(p^1, p^0, V_k, A_k),$$

and $s_k = M(p^0, V_k, A_k)/\Sigma M(p^0, V_k, A_k)$. This index can be interpreted as the relative cost of attaining a utility possibility

[72] The BLS produces several consumer price indexes. Unless otherwise stated, I am referring to the CPI-U.

[73] See Michael Boskin et. al. (1997).

[74] An excellent description of the CPI and the methods used to calculate it is provided by Brent Moulton (1996). See, also, Pollak (forthcoming).

[75] See Pollak (1983) for further discussion.

[76] Empirical implementation of household cost of living indexes are provided by Bert Balk (1990), Jorgenson and Slesnick (1983) and Mary Kokoski (1987).

[77] For a more exhaustive discussion of the aggregation problem in the context of the cost of living, see Pollak (1980, 1981).

frontier under the two sets of prices. As a measure of the cost of living, it has adverse ethical implications because the rich are more influential (i.e. have larger weight) than the poor in determining the cost of living.

To soften the normative basis, Prais also proposed a "democratic" index as the (unweighted) mean of the household indexes:

$$P^D(p^1, p^0, u) = (1/K)\Sigma P_k(p^1, p^0, V_k, A_k),$$

where K is the number of households. While Prais emphasized the normative implications of his approach, the welfare theoretic foundation for his index is, at best, ambiguous. Neither the plutocratic nor the democratic index answer the question of the amount of additional expenditure required to maintain a level of group welfare as prices change.

A solution to the conceptual problem of aggregation was provided by Pollak's (1981) social cost of living index. It is the ratio of the minimum cost of attaining social welfare W at current prices to that needed to attain the same welfare at base prices:

$$P(p^1, p^0, W) = \frac{M(p^1, W)}{M(p^0, W)},$$

where $M(\,)$ is a social expenditure function and W is the reference level of social welfare. The normative basis of this index is made explicit through the choice of a social welfare function. Most important, it can be used to compare well-being across groups of households as prices change. If the index doubles between 1980 and 1990, twice the expenditure is required to maintain the level of social welfare.

This added feature comes at the cost of adding a necessarily arbitrary element to the measurement of inflation because of the requirement of an explicit social welfare function. This suggests that aggregate price indexes are inherently normative and they cannot be expected to be invariant across different specifications of the social welfare function. Will one obtain radically different inflation estimates using a maximin social welfare function compared to utilitarian specification? For the United States, I've found that the underlying social welfare function has little influence on the price index (Slesnick 1991b). Pollak (forthcoming) points out, however, that the role of aggregation is likely to be more important (and is certainly more complex) once quality issues are introduced. That is, differential access to high quality goods will undoubtedly influence estimates of the average cost of living. Indeed, much recent discussion involves the role of substitution biases, the "new goods" problem and the role of quality changes on estimates of the cost of living. Scant little attention has been given to the interaction of these issues with the aggregation question.

In the empirical implementation of the aggregate cost of living, debate has centered on whether an econometric or index number approach should be used. Government agencies are reticent to use econometric models out of concern that widely-used inflation estimates will change when the model is reestimated or because the estimates are sensitive to model specification. Index number approaches, such as Diewert's superlative indexes, do not use an econometric model (although a functional form assumption on preferences is often implicit) but still require repeated observations on demands. Since these data are nonexistent at the household level, this method can only be implemented if demands are aggregated over groups of households and the averages are used to calculate the price index. This requires a representative agent assumption, with all the attendant problems, to provide a welfare theoretic interpretation.

Another practical concern is the level of aggregation at which the cost-of-living indexes should be applied. If the index is used to adjust benefits that are targeted to a subgroup of the population, it would seem appropriate to use a cost of living index that represents the consumption patterns of that group rather than a national price index. If social security benefits are to be indexed, a cost of living index specific to the elderly should be used. As Pollak (forthcoming) points out, the importance of this issue is ultimately an empirical issue and the evidence is scarce. Boskin and Michael Hurd (1985), Todd Idson and Cynthia Miller (1997) and Jorgenson and Slesnick (1998) find that inflation rates are similar across demographic groups while Robert Michael (1979) and Robert Hagemann (1982) report somewhat larger differences.

Given the fundamental importance of cost of living measurement to public policy, it is surprising that a number of potentially important issues have gone largely unnoticed in the empirical literature. The role of labor supply, public goods and intertemporal life cycle influences have not been addressed to the extent that they deserve, and all represent areas that would benefit from future research.

Inequality. Has everybody shared equally in the rise in the standard of living or have some segments of the population been left behind? Should growth be enhanced at the expense of a more egalitarian distribution? Is such a tradeoff inevitable or is it possible to have an increase in living standards without a rise in inequality? Fundamental questions such as these suggest that distributional issues, along with growth, are particularly influential in the design of public policy.

In quantifying the degree of dispersion in well-being, the policy analyst must decide on the representation of household welfare and the appropriate means of comparing distributions. Atkinson (1970) showed that ranking distributions is straightforward if the Lorenz curves do not intersect. If they do, then no unambiguous ordering is possible, and we must rely on summary statistics to quantify the levels of inequality. In this case, researchers must decide between a number of alternatives and we have no guarantee that judgements concerning levels and trends of inequality are invariant to this decision.[78] The choice of one index over another requires subjective judgements related to its sensitivity to transfers between households. The Gini coefficient is the ratio of the area between the Lorenz curve and the diagonal to the total area below the diagonal. It is most sensitive to changes in the middle of the distribution. The coefficient of variation is the ratio of the standard deviation to the mean, and mean-preserving changes in the distribution have the same effect no matter where they occur. Theil's entropy index is most sensitive to changes at the lower end of the distribution.[79]

The link between inequality indexes and their ethical basis can be made explicit by defining inequality to be the loss in social welfare attributable to disparities in welfare across households. This approach was pioneered by Atkinson (1970) who demonstrated that, for a fixed mean, Lorenz dominance of one distribution over another implies a higher level of social welfare for a utilitarian social welfare function regardless of the form of the household welfare

[78] The sensitivity of inequality estimates to the index used has been pointed out by Frank Levy and Richard Murnane (1992) in their study of earnings inequality in the U.S.

[79] For further discussion of the properties of a wide variety of inequality indexes, see Blackorby and Donaldson (1978) and Cowell (1977).

function.[80] Failing such a relationship, a functional form of the household welfare function must be specified. Assume, initially, that the welfare functions are identical across households and are functions only of income. Atkinson defined "equally distributed equivalent income" (say y^*) as the amount of income which, if allocated to everyone, yields the same level of social welfare as the observed distribution. It is implicitly represented as:

$$W(y^*, y^*, \ldots, y^*) = W(y_1, y_2, \ldots, y_K),$$

and the inequality index is defined as one minus the ratio of equally distributed equivalent income to mean income:

$$I(y_1, y_2, \ldots, y_K) = 1 - \frac{y^*}{\bar{y}}.$$

If household welfare (in this case, represented by income) is the same for each household, social welfare is maximized (for a fixed level of aggregate income), there is no inequality, and the index is equal to zero. If one household has all of the resources and the others nothing, social welfare takes its lowest possible value and the index is equal to one. Blackorby and Donaldson (1978) have shown that this general approach to inequality measurement has, as special cases, virtually all of the commonly used indexes including the Gini coefficient, Theil's entropy coefficient and the coefficient of variation.

While Atkinson's measure of inequality is not as commonly used as statistical measures of dispersion, it has several advantages. The normative basis of the index, which is part of any measure, is made explicit. This facilitates an examination of the sensitivity of conclusions to different assumptions about the magnitude by which inequality changes with mean-preserving transfers. Since the index is based on a social welfare function, the implied policy prescriptions are built in. For example, the gains that result from poverty relief are readily determined. As important, the inequality indexes are coherently related to the measures of social welfare.

Atkinson's framework can be extended to the case in which the welfare function depends on prices, expenditure and the demographic characteristics of the household. In particular, let the social welfare function W be defined as:

$$W = W(V(p, M_1, A_k), \ldots., V(p, M_K, A_K)),$$

where $V(\,)$ is an indirect utility function. If the social welfare function is equity-regarding, it is maximized at the perfectly egalitarian distribution of welfare and an inequality measure of the Atkinson type is given by:[81]

$$I(p, W, W_{\max}) = 1 - \frac{M(p, W)}{M(p, W_{\max})},$$

where $M(\,)$ is the social expenditure function, $W_{\max}$ is the social welfare attained at the perfectly egalitarian distribution, and W is the actual level of social welfare. As with Atkinson's original definition, this index lies between zero and one and is equal to zero when welfare is equalized across all households.

In recent empirical studies there is increasing concern with such aggregation issues, but the choice of the household welfare measure is almost an afterthought. The sheer volume of papers measuring income (or earnings) inequality suggests that a tremendous intellectual investment is associated with the use of annual income as a measure

[80] See, also, Serge Kolm (1969) and Sen (1973). Partha Dasgupta, Sen and David Starrett (1973) generalized Atkinson's (1970) result to include any social welfare function that is Schur-concave.

[81] This generalization was proposed by Jorgenson and Slesnick (1984) and subsequently used (Slesnick 1994) to examine trends in inequality in the United States.

of economic welfare.[82] The vast majority of these papers estimate inequality within a single period and, as I've argued repeatedly, income is unlikely to provide an accurate representation of welfare. The burning question is whether inequality measures that are founded on welfare functions with a sound theoretical foundation give noticeably different depictions of inequality.[83]

The biases associated with family or household income as a welfare measure arise from the use of income instead of expenditure, and the omission of price and demographic effects in measuring well-being. An implication of the permanent income hypothesis is that the distribution of total expenditure is quite different from the distribution of income. Households with low incomes are disproportionately represented by those with temporary reductions in current income and will have high ratios of consumption to income. Households with high income levels are over-represented by those with transitory increases in income and will exhibit low ratios of consumption to income. All other things equal, one would expect less dispersion in the distribution of total expenditure relative to the income distribution.

In fact, this is typically what is observed when the household income distribution is compared to the distribution of expenditures. Using data from the *Consumer Expenditure Surveys*, I've found (Slesnick 1994) that the expenditure distribution is more equally distributed than the income distribution in the U.S. and exhibits substantially less movement over time. Susan Mayer and Christopher Jencks (1992) and David Johnson and Stephanie Shipp (1997) also find the expenditure distribution to be more equally distributed compared to income but report that the trend varies with the sample definition, the consumption definition and the method of quantifying inequality. Regardless, with little doubt, the expenditure distribution differs significantly from the income distribution and its use biases inequality estimates in an important way.

The symmetric treatment of heterogeneous households, which is common in income inequality studies, implicitly assumes that households with the same income but different compositions are equally well-off. If household structure influences expenditure patterns, needs and welfare, such effects are likely to have an important impact on both the level and trend of inequality. In particular, the demographic changes that have taken place in the postwar United States suggest that movements in the family income distribution are likely to be quite different from those based on even the simplest form of an attribute-specific welfare measure.

The importance of household composition on inequality estimates can be seen by examining differences in household size at different points of the distribution. In the U.S., household expenditures and the average size of the household have a strong positive relationship. This suggests that adjustments for differences in household composition will have an important effect on the estimated level of inequality. Depending on the degree of economies of

[82] Reviews of the voluminous empirical literature are provided by Levy and Murnane (1992) for earnings inequality in the U.S. and Peter Gottschalk and Timothy Smeeding (1997) cross-nationally.

[83] Of greater interest are intertemporal or lifetime inequality estimates. These are few and far between although see Kathryn Shaw (1989) and Joel Slemrod (1992) and Fullerton and Rogers (1993). Deaton and Christina Paxson (1994) take an important step in actually linking intertemporal consumption behavior with observed patterns of within cohort inequality. This is one of the few papers that attempt to explain patterns of inequality rather than simply describe them.

scale in consumption, there will be a reordering of relative welfare levels, and a change in the estimated level of inequality. In fact, Coulter, Cowell and Jenkins (1992) find a systematic relationship between the scale economies and the level of inequality in the U.K. The relationship becomes more complicated once one accounts for other forms of heterogeneity (Banks and Preston 1994). In general, though, ignoring differences between households will yield biased estimates of inequality.[84]

A measure of welfare based solely on income also ignores the potential impact of prices on the distribution. Increases in the prices of necessities relative to luxuries will hurt the poor relatively more than the rich, and will thus increase the dispersion in welfare. The large increases in the relative prices of energy goods in the 1970s could have had a substantial impact on the relative welfare levels. This effect would obviously be missed by looking at the distribution of household income alone. Empirically, however, it appears that the distributional effects of price changes have been quite small.[85]

4. *Conclusions*

The influence of Harberger's (1971) proposal that consumer's surplus be used as a measure of individual and social welfare is reflected by its widespread use in applied welfare economics. Despite years of criticism, it remains the method of choice among nonspecialists. It is easy to use, the intuition underlying its interpretation as a welfare measure is transparent, and the data requirements for implementation are minimal.

Since Harberger's paper, the limitations and pitfalls of consumer's surplus have been demonstrated systematically and definitively by Chipman and Moore among many others. To avoid these conceptual problems, methods have been developed to estimate Hicksian surplus measures. These methods have the same data requirements as consumer's surplus and are just as easy to compute. Recent focus has been on the development of innovative methods of implementation that provide the greatest flexibility in modeling the welfare effects subject to the limitations imposed by the data.

While the conceptual issues of welfare measurement have been resolved at the micro level, there is more ambiguity as to how to proceed to aggregate the welfare effects of households. Most agree that the procedure of assuming a representative agent and applying the techniques designed for households to aggregate data has serious shortcomings. Moreover, the work of Blackorby and Donaldson, Roberts, and others shows that the use of income or other money-metric welfare indicators as arguments of a social welfare function also has many conceptual problems.

The rejection of these natural extensions of welfare economics to the aggregate level left a gap as to what to do in the critical area of measuring welfare for groups of households. Roberts and Sen provided a social choice theoretic framework by shifting the emphasis in the existing literature from the impossibility of the aggregation of welfare outcomes to conditions for possibility.

[84] Accounting for differences across households affects both the levels and trends of inequality in the U.S. The precise magnitudes of the differences depend on the characteristics used to distinguish households and the equivalence scales. See Slesnick (1994).

[85] The distributional effects of prices have been examined by Muellbauer (1974) and Newbery (1995) for the United Kingdom, Ravallion and Subramanian (1996) for India, and Stoker (1986) and Slesnick (1994) for the United States. Fullerton and Rogers (1997) find somewhat larger distributional effects when simulating tax policy in a computable general equilibrium model.

Their descriptions of the conditions necessary for the existence of social welfare functions facilitates consistent aggregation and a sound foundation for social policy evaluation, inequality and poverty measurement, and the measurement of the cost of living. Although the use of this framework is relatively new to practitioners, continuation of the recent research efforts will yield empirical methods of social welfare measurement that, like the micro level counter-parts, are based on a solid theoretical footing.

References

Abbott, Michael and Orley Ashenfelter. 1976. "Labor Supply, Commodity Demand and the Allocation of Time," *Rev. Econ. Stud.*, 43:135, pp. 389–411.

Afriat, S.N. 1967. "The Construction of a Utility Function from Expenditure Data," *Int. Econ. Rev.*, 8:1, pp. 67–77.

Ahmad, Ehtishan and Nicholas Stern. 1984. "The Theory of Reform and Indian Indirect Taxes," *J. Public Econ.*, 25:3, pp. 259–98.

Anand, Sudhir and Martin Ravallion. 1993. "Human Development in Poor Countries: On the Role of Private Incomes and Public Services," *J. Econ. Perspectives*, 7:1, pp. 133–50.

Anderson, Gordon and Richard Blundell. 1982. "Estimation and Hypothesis Testing in Dynamic Singular Equation Systems," *Econometrica*, 50:6, pp. 1559–71.

Apps, Patricia and Ray Rees. 1996. "Labor Supply, Household Production and Intra-Family Welfare Distribution," *J. Public Econ.*, 60:2, pp. 199–219.

——— and Elizabeth Savage. 1989. "Labour Supply, Welfare Rankings and the Measurement of Inequality," *J. Public Econ.*, 39:3, pp. 335–64.

Arrow, Kenneth J. 1951. *Social Choice and Individual Values.* NY: John Wiley & Sons.

Atkinson, Anthony B. 1970. "On Measurement of Inequality," *J. Econ. Theory*, 2:3, pp. 244–63.

———. 1992. "Measuring Poverty and Differences in Family Composition," *Economica*, 59:2, pp. 1–16.

——— and Francois Bourguignon. 1987. "Income Distribution and Differences in Needs," in *Arrow and the Foundation of the Theory of Economic Policy*. G.R. Feiwel, ed. London: Macmillan, pp. 350–70.

Auerbach, Alan J. 1985. "The Theory of Excess Burden and Optimal Taxation," in *Handbook of Public Economics*, Vol. 1. A.J. Auerbach and M. Feldstein, eds. Amsterdam: North Holland, pp. 61–127.

——— and Laurence Kotlikoff. 1987. *Dynamic Fiscal Policy*. Cambridge: Cambridge U. Press.

Balk, Bert M. 1990. "On Calculating Cost-of-Living Index Numbers for Arbitrary Income Levels," *Econometrica*, 58:1, pp. 75–92.

Banks, James; Richard Blundell, and Arthur Lewbel. 1995. "Quadratic Engel Curves, Indirect Tax Reform and Welfare Measurement," Institute Fiscal Stud. Work. Paper.

———. 1996. "Tax Reform and Welfare Measurement: Do We Need Demand System Estimation?," *Econ. J.*, 106:438, pp. 1227–41.

Banks, James; Richard Blundell, and Ian Preston. 1994. "Life-cycle Expenditure Allocations and the Consumption Costs of Children," *Europ. Econ. Rev.*, 38:7, pp. 1391–410.

——— and Paul Johnson. 1994. "Equivalence Scale Relativities Revisited," *Econ. J.*, 104:425, pp. 883–90.

Barnett, William A. 1979. "The Joint Allocation of Leisure and Goods," *Econometrica*, 47:3, pp. 539–63.

——— and Seungmook Choi. 1989. "A Monte Carlo Study of Tests of Blockwise Weak Separability," *J. Bus. Econ. Statist.*, 7:3, pp. 363–77.

Barnett, William A. and Yul Lee. 1985. "Global Properties of the Minflex Laurent, Generalized Leontief, and Translog Flexible Functional Forms," 53:6, pp. 1421–37.

Barten, A.P. 1964. "Family Composition, Prices, and Expenditure Patterns," in *Econometric Analysis for National Economic Planning: 16th Symposium of the Colston Society*. P. Hart, G. Mills and J.K. Whitaker, eds. London: Butterworth, pp. 277–92.

Becker, Gary S. 1981. *A Treatise on the Family*. Cambridge: Harvard U. Press.

———, Michael Grossman, and Kevin Murphy. 1994. "An Empirical Analysis of Cigarette Addiction," *Amer. Econ. Rev.*, 84:3, pp. 396–418.

Bergson, Abram. 1938. "A Reformulation of Certain Aspects of Welfare Economics," *Quart. J. Econ.*, 52:1, pp. 310–34.

Bierens, Herman J. and Pott-Buter, Hettie A. 1990. "Specification of Household Engel Curves by Nonparametric Regression," *Econometric Rev.*, 9:2, pp. 123–84.

Blackorby, Charles and David Donaldson. 1978. "Measures of Relative Inequality and their Meaning in Terms of Social Welfare," *J. Econ. Theory*, 18:1, pp. 651–75.

———. 1985. "Consumers' Surpluses and Consistent Cost-Benefit Tests," *Social Choice and Welfare*, 1, pp. 251–62.

———. 1988a. "Money Metric Utility: A Harmless Normalization?" *J. Econ. Theory*, 46:1, pp. 120–29.

———. 1988b. "Adult-Equivalence Scales and the Economic Implementation of Interpersonal Comparisons of Well-Being," Disc. Paper 88–27, U. British Columbia.

———. 1990. "A Review Article: The Case against the Use of the Sum of Compensating Variations in Cost–Benefit Analysis," *Can. J. Econ.*, 23:3, pp. 471–94.

——— and David Moloney. 1984. "Consumer's Surplus and Welfare Change in a Simple Dynamic Model," *Rev. Econ. Stud.*, 51:164, pp. 171–76.

Blackorby, Charles; Daniel Primont, and Robert Russell. 1978. *Duality, Separability, and Functional Structure: Theory and Economic Applications*. NY: North-Holland.

Blundell, Richard. 1988. "Consumer Behavior: Theory and Empirical Evidence—A Survey," *Econ. J.*, 98:389, pp. 16–65.

——— and Arthur Lewbel. 1991. "The Information Content of Equivalence Scales," *J. Econometrics*, 50:1/2, pp. 49–68.

Blundell, Richard; Martin Browning, and Costas Meghir. 1994. "Consumer Demand and the Life Cycle Allocation of Household Expenditure," *Rev. Econ. Stud.*, 61:1, pp. 57–81.

Blundell, Richard; Panos Pashardes, and Guglielmo Weber. 1993. "What Do We Learn about Demand Patterns from Micro Data?," *Amer. Econ. Rev.*, 83:3, pp. 570–97.

Blundell, Richard, et. al. 1988. "Labour Supply Specification and the Evaluation of Tax Reforms," *J. Public Econ.*, 36:1, pp. 23–52.

Boskin, Michael, et. al. 1997. "The CPI Commission: Findings and Recommendations," *Amer. Econ. Rev.*, 87:2, pp. 78–83.

Boskin, Michael and Michael Hurd. 1985. "Indexing Social Security Benefits: A Separate Price Index for the Elderly," *Public Finance Quart.*, 13:4, pp. 436–49.

Browning, Martin. 1992. "Children and Household Economic Behavior," *J. Econ. Lit.*, 30:3, pp. 1434–75.

——— and Costas Meghir. 1991. "The Effects of Male and Female Labor Supply on Commodity Demands," *Econometrica*, 59:4, pp. 925–51.

Browning, Martin et. al. 1994. "Income and Outcomes: A Structural Model of Intrahousehold Allocation," *J. Polit. Econ.*, 102:6, pp. 1067–96.

Buhmann, Brigitte, et. al. 1988. "Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study Database, *Rev. Income Wealth*, 34:2, pp. 115–42.

Byron, R.P. and Anil Bera. 1983. "Least Squares Approximations to Unknown Regression Functions: A Comment," *Int. Econ. Rev.*, 24:1, pp. 255–60.

Casperson, Erik and Gilbert Metcalf. 1994. "Is the Value Added Tax Regressive? Annual Versus Lifetime Incidence Measures," *Nat. Tax J.*, 47:4, pp. 731–46.

Chiappori, Pierre A. 1988. "Rational Household Labor Supply," *Econometrica*, 56:1, pp. 63–89.

———. 1992. "Collective Labor Supply and Welfare," *J. Polit. Econ.*, 100:3, pp. 437–67.

Chipman, John S. and James Moore. 1971. "The Compensation Principle in Welfare Economics," in *Papers in Quantitative Economics*. A. Zarley and J. Moore, eds. Lawrence: U. Press of Kansas, pp. 1–77.

———. 1973. "Aggregate Demand, Real National Income, and the Compensation Principle," *Int. Econ. Rev.*, 14:1, pp. 153–81.

———. 1976a. "Why an Increase in GNP Need Not Imply an Improvement in Potential Welfare," *Kyklos*, 29:3, pp. 391–418.

———. 1976b. "The Scope of Consumer's Surplus Arguments," in *Evolution, Welfare and Time in Economics: Essays in Honor of Nicholas Georgescu–Roegen*. A.M. Tang et. al., eds. Lexington: Heath-Lexington Books, pp. 69–123.

———. 1980. "Compensating Variation, Consumer's Surplus, and Welfare," *Amer. Econ. Rev.*, 70:5, pp. 933–49.

Christensen, Laurits; Dale Jorgenson, and Lawrence Lau. 1975. "Transcendental Logarithmic Utility Functions," *Amer. Econ. Rev.*, 65:3, pp. 367–83.

Cooter, Robert and Peter Rappoport. 1984. "Were the Ordinalists Wrong About Welfare Economics?" *J. Econ. Lit.*, 22:2, pp. 507–30.

Coulter, Fiona; Frank Cowell, and Jenkins, Stephen. 1992. "Equivalence Scale Relativities and the Extent of Inequality and Poverty," *Econ. J.*, 102:414, pp. 1067–82.

Cowell, Frank A. 1977. *Measuring Inequality*. NY: Halsted Press.

Creel, Michael D. 1997. "Welfare Estimation Using Fourier Form: Simulation Evidence for the Recreation Demand Case," *Rev. Econ. Statist.*, 79:1, pp. 88–94.

Dasgupta, Partha; Amartya Sen, and David Starrett. 1973. "Notes on the Measurement of Inequality," *J. Econ. Theory*, 6:2, pp. 180–87.

Deaton, Angus. 1977. "Equity, Efficiency, and the Structure of Indirect Taxation," *J. Public Econ.*, 8:3, pp. 299–312.

———. 1979. "The Distance Function in Consumer Behaviour with Applications to Index Numbers and Optimal Taxation," *Rev. Econ. Stud.*, 46:144, pp. 391–405.

———. 1989. "Rice Prices and Income Distribution in Thailand: A Nonparametric Analysis," *Econ. J.*, 99:395, pp. 1–37.

——— and John Muellbauer. 1980a. *Economics and Consumer Behavior.* Cambridge: Cambridge U. Press.

———. 1980b. "An Almost Ideal Demand System," *Amer. Econ. Rev.*, 1980b, 70:3, pp. 312–26.

———. 1986. "On Measuring Child Costs: With Applications to Poor Countries," *J. Polit. Econ.*, 94:4, pp. 720–44.

Deaton, Angus and Christina Paxson. 1994. "Intertemporal Choice and Inequality," *J. Polit. Econ.*, 3:102, pp. 437–67.

De Borger, Bruno. 1989. "Estimating the Welfare Implications of In-Kind Government Programs," *J. Public Econ.*, 38:2, pp. 215–26.

Diamond, Peter A. and Daniel L. McFadden. 1974. "Some Uses of the Expenditure Function in Public Finance," *J. Public Econ.*, 3:1, pp. 3–21.

Diewert, W. Erwin. 1971. "An Application of the Shepherd Duality Theorem: A Generalized

Leontief Production Function," *J. Polit. Econ.*, 79:3, pp. 481–507.

———. 1974. "Applications of Duality Theory," in *Frontiers of Quantitative Economics,* Vol. 2. M.D. Intriligator and D.A. Kendrick, eds. Amsterdam: North-Holland.

———. 1976a. "Exact and Superlative Index Numbers," *J. Econometrics*, 4:2, pp. 115–45.

———. 1976b. "Harberger's Welfare Indicator and Revealed Preference Theory," *Amer. Econ. Rev.*, 66:1, pp. 143–52.

———. 1981. "The Econ. Theory of Index Numbers: A Survey," in *Essays in the Theory and Measurement of Consumer Behavior.* A. Deaton, ed. Cambridge, Cambridge U. Press, pp. 163–208.

———. 1990. "The Theory of the Cost-of-Living Index and the Measurement of Welfare Change" in *Price Level Measurement*. W.E. Diewert, ed. Amsterdam, North-Holland, pp. 79–147.

———. 1992. "Exact and Superlative Welfare Change Indicators," *Econ. Inquiry*, 30:4, pp. 565–83.

——— and Terence Wales. 1987. "Flexible Functional Forms and Global Curvature Conditions," *Econometrica*, 55:1, pp. 43–68.

———. 1995. "Flexible Functional Forms and Tests of Homogeneity," *J. Econometrics*, 67:2, pp. 259–302.

Elbadawi, Ibrahim; A. Ronald Gallant, and Geraldo Souza. 1983. "An Elasticity Can Be Estimated Consistently without A Priori Knowledge of Functional Form," *Econometrica*, 51:6, pp. 1731–51.

Engel, Ernst. 1895. "Die Lebenskosten Belgischer Arbeiter–Familien Fruher und Jetzt," *Int. Statist. Institute Bull.*, 9, pp. 1–74.

Famulari, Melissa. 1995. "A Household-Based, Nonparametric Test of Demand Theory," *Rev. Econ. Statist.*, 77:2, pp. 371–82.

Fortin, Bernard and Guy Lacroix. 1997. "A Test of the Unitary and Collective Models of Household Labour Supply," *Econ. J.*, 107:443, pp. 933–55.

Fullerton, Don and Diane Rogers. 1993. *Who Bears the Lifetime Tax Burden?* Washington: Brookings Institution.

———. 1997. "Neglected Effects on the Uses Side: Even a Uniform Tax Would Change Relative Goods Prices," *Amer. Econ. Rev.*, 87:2, pp. 120–25.

Gallant, A. Ronald. 1981. "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Functional Form," *J. Econometrics*, 15:2, pp. 211–45.

Gorman, William M. 1953. "Community Preference Fields," *Econometrica*, 21:1, pp. 63–80.

———. 1955. "The Intransitivity of Certain Criteria used in Welfare Economics," *Oxford Econ. Pap.*, 7, pp. 25–35.

———. 1959. "Separable Utility and Aggregation," *Econometrica*, 27:3, pp. 469–81.

———. 1976. "Tricks with Utility Functions," in *Essays in Economic Analysis*. M. Artis and M. Nobay, eds. Cambridge: Cambridge U. Press, pp. 211–43.

———. 1981. "Some Engel Curves," in *Essays in the Theory and Measurement of Consumer Behavior*. A. Deaton, ed. Cambridge: Cambridge U. Press.

Gottschalk, Peter and Timothy Smeeding. 1997. "Cross-National Comparisons of Earnings and Income Inequality," *J. Econ. Lit.*, 35:2, pp. 633–87.

Haddad, Lawrence and Ravi Kanbur. 1990. "How Serious is the Neglect of Intra-Household Inequality," *Econ. J.*, 100:3, pp. 866–81.

Hagemann, Robert P. 1982. "The Variability of Inflation Rates across Household Types," *J. Money, Credit, Banking*, 14:1, pp. 494–510.

Hammond, Peter J. 1976. "Equity, Arrow's Conditions, and Rawls' Difference Principle," *Econometrica*, 44:4, pp. 793–804.

———. 1977. "Dual Interpersonal Comparisons of Utility and the Welfare Economics of Income Distribution," *J. Public Econ.*, 7:1, pp. 51–71.

Hanemann, Michael and Edward Morey. 1992. "Separability, Partial Demand Systems, and Consumer's Surplus Measures," *J. Environ. Econ. Manage.*, 22:3, pp. 241–58.

Harberger, Arnold C. 1971. "Three Basic Postulates for Applied Welfare Economics," *J. Econ. Lit.*, 9:3, pp. 785–97.

Hardle, Wolfgang. 1990. *Applied Nonparametric Regression*. Cambridge: Cambridge U. Press.

———, Werner Hildenbrand, and Michael Jerison. 1991. "Empirical Evidence on the Law of Demand," *Econometrica*, 59:6, pp. 1525–49.

Hausman, Jerry A. 1981. "Exact Consumer's Surplus and Deadweight Loss," *Amer. Econ. Rev.*, 71:4, pp. 662–76.

———. 1985. "The Econometrics of Nonlinear Budget Sets," *Econometrica*, 53:6, pp. 1255–82.

——— and Whitney Newey. 1995. "Nonparametric Estimation of Exact Consumers Surplus and Deadweight Loss," *Econometrica*, 63:6, pp. 1445–76.

——— and Paul Ruud. 1984. "Family Labor Supply with Taxes," *Amer. Econ. Rev.*, 74:2, pp. 242–48.

Haveman, Robert; Mary Gabay, and James Andreoni. 1987. "Exact Consumer's Surplus and Deadweight Loss: A Correction," *Amer. Econ. Rev.*, 77:3, pp. 494–95.

Hicks, John R. 1942. "Consumer's Surplus and Index Numbers," *Rev. Econ. Stud.*, 9:2, pp. 126–37.

Houthakker, Hendrik S. 1957. "An Int. Comparison of Household Expenditure Patterns Commemorating the Centenary of Engel's Law," *Econometrica*, 25:4, pp. 532–51.

Hudson, Edward A. and Dale Jorgenson. 1974. "U.S. Energy Policy and Economic Growth, 1975–2000," *Bell J. Econ. Manage. Sci.*, 5:2, pp. 461–514.

Hurwicz, Leonid and Hirofumi Uzawa. 1971. "On the Integrability of Demand Functions," in

Preferences, Utility and Demand. J. Chipman et. al. eds. NY: Harcourt, Brace Jovanovich, pp. 114–48.

Idson, Todd and Cynthia Miller. 1997. "The Implications of Demographic-Specific Inflation Rates for Trends in Real Educational Wage Differentials," *J. Bus. Econ. Statist*., 15:4, pp. 464–69.

Johnson, David S. and Stephanie Shipp. 1997. "Trends in Consumption Inequality in the United States," *Rev. Income Wealth*, 43:2, pp. 133–52.

Jorgenson, Dale W.; Lawrence Lau, and Thomas Stoker. 1980. "Welfare Comparison and Exact Aggregation," *Amer. Econ. Rev*., 70:2, pp. 268–72.

———. 1982. "The Transcendental Logarithmic Model of Aggregate Consumer Behavior," in *Advances in Econometrics*. Vol. 1. R.L. Basmann and G.F. Rhodes, eds. Greenwich: JAI Press, pp. 97–238.

Jorgenson, Dale W. and Daniel T. Slesnick. 1983. "Individual and Social Cost of Living Indexes," in *Price Level Measurement*. W.E. Diewert and C. Montmarquette, eds. Ottawa: Statistics Canada, pp. 241–323.

———. 1984. "Aggregate Consumer Behavior and the Measurement of Inequality," *Rev. Econ. Stud*., 51:166, pp. 369–92.

———. 1985a. "Efficiency vs Equity in Petroleum Taxation," *Energy J*., 6, pp. 171–88.

———. 1985b. "Efficiency versus Equity in Natural Gas Price Regulation," *J. Econometrics*, 30:1/2, pp. 301–16.

———. 1987. "Aggregate Consumer Behavior and Household Equivalence Scales," *J. Bus. Econ. Statist*., 5:2, pp. 219–32.

———. 1998. "Indexing Government Programs for Changes in the Cost of Living," unpublished manuscript.

——— and Peter Wilcoxen. 1991. "Carbon Taxes and Economic Welfare," *Brookings Pap. Econ. Act.: Microeconomics*, pp. 393–454.

Jorgenson, Dale W. and Kun-Young Yun. 1990. "Tax Reform and U.S. Economic Growth," *J. Polit. Econ*., 98:5, pp. s151–s193.

Kay, J.A. 1980. "The Deadweight Loss from a Tax System," *J. Public Econ*., 13:1, pp. 111–19.

Keen, Michael. 1990. "Welfare Analysis and Intertemporal Substitution," *J. Public Econ*., 42:1, pp. 47–66.

King, Mervyn A. 1983. "Welfare Analysis of Tax Reforms Using Household Data," *J. Public Econ*., 21:2, pp. 183–214.

Kirman, Alan P. 1992. "Whom or What Does the Representative Individual Represent?," *J. Econ. Perspectives*, 6:2, pp. 117–36.

Kokoski, Mary F. 1987. "Problems in the Measurement of Cost-of-Living Indexes," *J. Bus. Econ. Statist.,* 5:1, pp. 39–46.

Kolm, Serge C. 1969. "The Optimal Production of Social Justice," in *Public Economics*. J. Margolis and H. Guitton, eds. London: Macmillan, pp. 145–200.

Konus, A.A. 1939. "The Problem of the True Index of the Cost-of-Living," *Econometrica*, 7:1, pp. 10–29.

LaFrance, Jeffrey. 1993. "Weak Separability in Applied Welfare Analysis," *Amer. J. Agr. Econ*., 75:3, pp. 770–75.

Lau, Lawrence. 1982. "A Note on the Fundamental Theorem of Exact Aggregation," *Econ. Letters*, 9:2, pp. 119–26.

Levy, Frank and Richard Murnane. 1992. "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations," *J. Econ. Lit*., 30:3, pp. 1333–81.

Lewbel, Arthur. 1985. "A Unified Approach to Incorporating Demographic or other Effects into Demand Systems," *Rev. Econ. Stud*., 52:168, pp. 1–18.

———. 1989a. "Household Equivalence Scales and Welfare Comparisons," *J. Public Econ*., 39:3, pp. 377–91.

———. 1989b. "Exact Aggregation and a Representative Consumer," *Quart. J. Econ*., 104:3, pp. 622–33.

———. 1989c. "A Demand System Rank Theorem," *Econometrica*, 57:3, pp. 701–705.

———. 1990. "Full Rank Demand Systems," *Int. Econ. Rev*., 31:2, pp. 289–300.

———. 1991. "The Rank of Demand Systems: Theory and Nonparametric Estimation," *Econometrica*, 59:3, pp. 711–30.

———. 1995. "Consistent Nonparametric Hypothesis Tests with an Application to Slutsky Symmetry," *J. Econometrics*, 67:2, pp. 379–401.

———. 1996. "Aggregation without Separability: A Generalized Composite Commodity Theorem," *Amer. Econ. Rev*., 86:3, pp. 524–43.

———. Forthcoming "Consumer Demand Systems and Household Equivalence Scales," *Handbook of Applied Econometrics*.

——— and Richard Weckstein. Forthcoming. "Equivalence Scales, Cost of Children and Wrongful Death Laws," *J. Income Distribution*.

Lundberg, Shelly. 1988. "Labor Supply of Husbands and Wives: A Simultaneous Equations Approach," *Rev. Econ. Statist*., 70:2, pp. 224–35.

——— and Robert Pollak. 1993. "Separate Spheres Bargaining and the Marriage Market," *J. Polit. Econ*., 101:6, pp. 988–1010.

———. 1996. "Bargaining and Distribution in Marriage," *J. Econ. Perspectives*, 10:4, pp. 139–58.

McElroy, Marjorie and Mary Horney. 1981. "Nash Bargained Household Decisions: Toward a Generalization of the Theory of Demand," *Int. Econ. Rev*., 22:2, pp. 333–49.

McKenzie, Lionel. 1957. "Demand Theory without a Utility Index," *Rev. Econ. Stud*., 24:65, pp. 185–89.

McKenzie, George W. 1979. "Consumer's Surplus without Apology: Comment," *Amer. Econ. Rev*., 69:3, pp. 465–68.

——— and Ivor Pearce. 1976. "Exact Measures of Welfare and the Cost of Living," *Rev. Econ. Stud*., 43:3, pp. 465–68.

Malmquist, S. 1953. "Index Numbers and Indifference Surfaces," *Trabajos de Estadistica*, 4, pp. 209–42.

Manser, Marilyn. 1987. "Cash-Equivalent Values from In-Kind Benefits: Estimates from a Complete Demand System Using Household Data," BLS Work. Paper 173.

——— and Murray Brown. 1980. "Marriage and Household Decision-Making: A Bargaining Analysis," *Int. Econ. Rev.*, 21:1, pp. 31–44.

——— and Richard McDonald. 1988. "An Analysis of Substitution Bias in Measuring Inflation, 1959–1985," *Econometrica*, 56:4, pp. 909–30.

Maskin, Eric. 1978. "A Theorem on Utilitarianism," *Rev. Econ. Stud.*, 45:139, pp. 93–96.

Mayer, Susan and Christopher Jencks. 1992. "Recent Trends in Economic Inequality in the United States: Income vs. Expenditures vs. Material Well-Being," unpublished manuscript.

Mayshar, Joram and Shlomo Yitzhaki. 1995. "Dalton-Improving Indirect Tax Reform," *Amer. Econ. Rev.*, 85:4, pp. 793–807.

Meghir, Costas and Guglielmo Weber. 1996. "Intertemporal Nonseparability or Borrowing Restrictions? A Disaggregate Analysis Using a U.S. Consumption Panel," *Econometrica*, 64:5, pp. 1151–81.

Michael, Robert T. 1979. "Variation across Households in the Rate of Inflation," *J. Money, Credit, Banking*, 11:1, pp. 32–46.

Mohring, Herbert. 1971. "Alternative Welfare Gain and Loss Measures," *Western Econ. J.*, 9, pp. 349–68.

Moulton, Brent. 1996. "Bias in the Consumer Price Index: What is the Evidence?," *J. Econ. Perspectives*, 10:4, pp. 159–77.

Muellbauer, John. 1974. "Prices and Inequality: The U.K. Experience," *Econ. J.*, 84:333, pp. 32–55.

———. 1976. "Community Preferences and the Representative Consumer," *Econometrica*, 44:5, pp. 711–30.

———. 1977. "Testing the Barten Model of Household Composition Effects and the Cost of Children," *Econ. J.*, 87:347, pp. 460–87.

———. 1980. "The Estimation of the Prais-Houthakker Model of Equivalence Scales," *Econometrica*, 48:1, pp. 153–76.

Neary, J.P. and Kevin Roberts. 1980. "The Theory of Household Behavior under Rationing," *Europ. Econ. Rev*, 13:1, pp. 25–42.

Newbery, David W. 1995. "The Distributional Impact of Price Changes in Hungary and the United Kingdom," *Econ. J.*, 105:431, pp. 847–63.

Orshansky, Mollie. 1965. "Counting the Poor: Another Look at the Poverty Profile," *Soc. Sec. Bull.*, 28:1, pp. 3–29.

Pashardes, Panos. 1991. "Contemporaneous and Intertemporal Child Costs," *J. Public Econ.*, 45:2, pp. 191–213.

Pollak, Robert A. 1980. "Group Cost of Living Indexes," *Amer. Econ. Rev.*, 70:2, pp. 273–78.

———. 1981. "The Social Cost of Living Index," *J. Public Econ.*, 15:3, pp. 311–36.

———. 1983. "The Theory of the Cost-of-Living Index" in *Price Level Measurement.* W.E. Diewert and C. Montmarquette, eds. Ottawa: Statistics Canada, pp. 241–323.

———. 1991. "Welfare Comparisons and Situation Comparisons," *J. Econometrics,* 50:1/2, pp. 31–48.

———. Forthcoming. "The Consumer Price Index: A Research Agenda and Three Proposals," *J. Econ. Perspectives*.

——— and Terence Wales. 1979. "Welfare Comparisons and Equivalent Scales," *Amer. Econ. Rev.*, 69:2, pp. 216–21.

———. 1981. "Demographic Variables in Demand Analysis," *Econometrica*, 1981, 49:6, pp. 1533–51.

———. 1992. *Demand System Specification and Estimation*. Oxford: Oxford U. Press.

Poterba, James. 1989. "Lifetime Incidence and the Distributional Burden of Excise Taxes," *Amer. Econ. Rev.,* 79:2, pp. 325–30.

———. 1991. "Is the Gasoline Tax Regressive?," in *Tax Policy and the Economy*. D. Bradford, ed. Vol. 5, Cambridge: MIT Press.

Prais, S.J. 1959. "Whose Cost of Living?" *Rev. Econ. Stud.*, 26:70, pp. 126–34.

——— and Hendrik S. Houthakker. 1955. *The Analysis of Family Budgets.* Cambridge: Cambridge U. Press.

Ransom, Michael. 1987. "An Empirical Model of Discrete and Continuous Choice in Family Labor Supply," *Rev. Econ. Statist.,* 69:3, pp. 465–72.

Ravallion, Martin and S. Subramanian. 1996. "Welfare Measurement with and without Substitution," World Bank unpublished manuscript.

Robbins, Lionel. 1932. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.

———. 1938. "Interpersonal Comparisons of Utility: A Comment," *Econ. J.*, 43:192, pp. 635–41.

Roberts, Kevin W.S. 1980a. "Price Independent Welfare Prescriptions," *J. Public Econ.*, 13:3, pp. 277–98.

———. 1980b. "Possibility Theorems with Interpersonally Comparable Welfare Levels," *Rev. Econ. Stud.*, 47:147, pp. 409–20.

———. 1980c. "Interpersonal Comparability and Social Choice Theory," *Rev. Econ. Stud.*, 47:147, pp. 421–39.

———. 1980d. "Social Choice Theory: The Single Profile and Multi–Profile Approaches," *Rev. Econ. Stud.*, 47:147, pp. 441–50.

Rosen, Harvey. 1978. "The Measurement of Excess Burden with Explicit Utility Functions," *J. Polit. Econ.*, 86:2, pp. S121–36.

Ruiz-Castillo, Javier. 1987. "Potential Welfare and the Sum of the Individual Compensating or Equivalent Variations," *J. Econ. Theory*, 41:1, pp. 34–53.

Samuelson, Paul A. 1947. *Foundations of Economic Analysis*. Cambridge: Harvard U. Press.

———. 1948. "Consumption Theory in Terms of Revealed Preference" *Economica*, 15:60, pp.

243–53.
———. 1956. "Social Indifference Curves," *Quart. J. Econ.*, 70:1, pp. 1–22.
Schultz, T. Paul. 1990. "Testing the Neoclassical Model of Family Labor Supply and Fertility," *J. Human Res.*, 25:3, pp. 599–634.
Schwab, Robert. 1985. "The Benefits of In-Kind Government Programs," *J. Public Econ.*, 27:2, pp. 195–210.
Sen, Amartya K. 1970. *Collective Choice and Social Welfare*. San Francisco: Holden Day.
———. 1973. *On Economic Inequality*. Oxford: Clarendon Press.
———. 1977. "On Weights and Measures: Informational Constraints in Social Welfare Analysis," *Econometrica*, 45:7, pp. 1539–71.
———. 1979. "The Welfare Basis of Real Income Comparisons: A Survey," *J. Econ. Lit.*, 17:1, pp. 1–45.
———. 1985. *Commodities and Capabilities*. Amsterdam: North Holland.
———. 1987. *The Standard of Living*. Cambridge: Cambridge U. Press.
———. 1995. "Rationality and Social Choice," *Amer. Econ. Rev.*, 85:1, pp. 1–24.
Shaw, Kathryn L. 1989. "Intertemporal Labor Supply and the Distribution of Family Labor Supply," *Rev. Econ. Statist.*, 71:2, pp. 196–205.
Shoven, John and John Whalley. 1972. "A General Equilibrium Calculation of the Effects of Differential Taxation of Income from Capital in the U.S.," *J. Public Econ.*, 1:3/4, pp. 281–321.
Silberberg, Eugene. 1972. "Duality and the Many Consumer's Surpluses," *Amer. Econ. Rev.*, 62:5, pp. 942–52.
Slemrod, Joel. 1992. "Taxation and Inequality: A Time-Exposure Perspective," Work. Paper No. 3999, National Bureau Econ. Research.
Slesnick, Daniel T. 1991a. "Aggregate Deadweight Loss and Money Metric Social Welfare," *Int. Econ. Rev.*, 32:1, pp. 123–46.
———. 1991b. "Normative Index Numbers," *J. Econometrics*, 50:1/2, pp. 107–30.
———. 1992. "Aggregate Consumption and Saving in the Postwar United States," *Rev. Econ. Statist.*, 74:4, pp. 585–97.
———. 1994. "Consumption, Needs and Inequality," *Int. Econ. Rev.*, 35:3, pp. 677–703.
———. 1996. "Consumption and Poverty: How Effective are In-Kind Transfers?" *Econ. J.*, 106:439, pp. 1527–45.
Slivinski, Alan. 1983. "Income Distribution Evaluation and the Law of One Price," *J. Public Econ.*, 20:1, pp. 103–12.
Sonnenschein, Hugo. 1972. "Market Excess Demand Functions," *Econometrica*, 40:3, pp. 549–63.
Stoker, Thomas. 1986. "The Distributional Welfare Effects of Rising Prices in the United States," *Amer. Econ. Rev.*, 76:3, pp. 335–49.
———. 1993. "Empirical Approaches to the Problem of Aggregation over Individuals," *J. Econ. Lit.*, 31:4, pp. 1827–74.
Stone, Richard. 1954. "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand," *Econ. J.*, 64:255, pp. 511–27.
Taylor, Angus E. and Robert Mann. 1972. *Advanced Calculus*, 2nd ed., Lexington: Xerox College Publishing.
Thomas, Duncan. 1990. "Intra-Household Resource Allocation: An Inferential Approach," *J. Human Res.*, 25:4, pp. 635–64.
———. 1994. "Like Father, Like Son: Like Mother, Like Daughter: Parental Resources and Child Height," *J. Human Res.*, 29:4, pp. 950–88.
Tornqvist, Leo. 1936. "The Bank of Finland's Consumption Price Index," *Bank Finland Monthly Bull.*, 10, pp. 1–8.
Varian, Hal R. 1982. "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50:4, pp. 945–73.
Vartia, Yrjo O. 1983. "Efficient Methods of Measuring Welfare Change and Compensated Income in Terms of Ordinary Demand Functions," *Econometrica*, 51:1, pp. 79–98.
Weitzman, Martin. 1988. "Consumer's Surplus as an Exact Approximation when Prices are Appropriately Deflated," *Quart. J. Econ.*, 103:3, pp. 543–54.
White, Halbert. 1980. "Using Least Squares to Approximate Unknown Regression Functions," *Int. Econ. Rev.*, 21:1, pp. 149–70.
Willig, Robert E. 1976. "Consumer's Surplus without Apology," *Amer. Econ. Rev.*, 66:4, pp. 589–97.
Wolak, Frank A. 1996. "The Welfare Impacts of Competitive Telecommunications Supply: A Household-Level Analysis," *Brookings Papers Econ. Act.: Microeconomics*, pp. 269–340.

Empirical-Quantitative Approaches ii. Acknowledgments Abstract. Table of Contents.Â Empirical-Quantitative Approaches 1. For understandable reasons, case selection in most studies [on international regimes] has been driven by practical considerations instead of methodological requirements. Moreover, the choice of both dependent and independent variables for systematic attention in these small-n case studies has failed, in general, to produce a cumulative and consistent set of information on an agreed upon set of important variables.Â The measurement of any phenomenon always contains a certain element of chance error. The goal of error-free measurement - while laudable - is never attained in any area of scientific investigation (Carmines & Zeller, 1986, 11, emphasis in the original).