SOCIAL INDICATORS AND TRUE COMPARISONS OF LIVING STANDARDS

Steve Dowrick, Yvonne Dunlop and John Quiggin

Australian National University

28th March 2001

Abstract

The construction of an international index of standards of living which incorporates social indicators as well as economic output typically involves scaling and weighting procedures which have no foundation in welfare theory. Moreover, elements of GDP which are intermediate inputs into the production of household welfare are often misrepresented as final consumption goods. Revealed preference axioms can be used to make quality-of-life comparisons if it is possible to estimate the representative household's production technology for the social indicators. This method is applied to make comparisons over the components of GDP and life expectancy for a cross-section of 58 countries. Neither GDP rankings, nor the rankings of the Human Development Index, are consistent with the partial ordering of revealed preference. A method of constructing a utility-consistent index incorporating both consumption and life expectancy is suggested.

JEL: I31, O47

Communications to: John Quiggin, Faculty of Economics, Copland Building (no. 26),

Australian National University, ACT 0200, Australia. Tel: +61 2 6125 4606;

e-mail: John.Quiggin @anu. edu. au

We thank Nancy Wallace for helpful comments and criticism

SOCIAL INDICATORS AND TRUE COMPARISONS OF LIVING STANDARDS

International comparisons of living standards and development are inescapable. Public opinion, inasmuch as it is represented by the media, has an insatiable appetite for world or regional rankings. Changes in these rankings are regularly used by politicians and policy analysts as a basis for assessing the efficacy of national policies. Governments and international organizations use measures of development to allocate funding to countries or regions. [1] The underlying indexes of national progress are used extensively in econometric testing of theories of growth and development. Here we propose a contribution to the method by which such international comparisons are made, based on the economic theory of revealed preference.

Most commonly, international comparisons of living standards or development are made in terms of gross domestic product (GDP) per capita. Such comparisons have been criticized on a number of grounds. The most common criticism is based on the fact that standard GDP indexes are more properly regarded as partial measures of aggregate output than as indicators of well-being. No allowance is made for environmental differences, non-market household production, or production and consumption externalities. For example, if polluting industries cause illnesses that require expensive medical treatment, both the output of the polluting industry and the expenditure on medical services will be counted as part of GDP. Thus, aggregate output may increase although well-being, as measured by various social indicators, declines.

Here we develop a method of welfare comparison that incorporates both GDP and social indicators. Our approach treats the average outcome in a country, in terms of real consumption and life expectancy, as the choice of representative agents facing the relative

[1] For instance, the European Community has allocated EUR213 billion under the European Regional Development Fund to regions where GDP per person is less than 75 per cent of the Community average (http://www. europa. eu. int/comm/regional_policy/activity/erdf/erd1b_en. htm).

prices and budget constraints of each country. Following Kravis, Heston and Summers (1982), revealed preference may then be used to compare per capita bundles and to test the hypothesis of common tastes amongst representative agents.

The paper is organized as follows. In Section I, we review the literature on extended measures of GDP, designed to take account of variables such as life expectancy. In Section II we briefly summarize the relevant theory of revealed preference.

In Section III, we extend standard GDP comparisons by incorporating evidence on life expectancy across 58 countries. In our empirical application, we have chosen to concentrate on length of life as the single most important of the social indicators in common use. Marketed goods such as food, medical services and education are important determinants of mortality, so we are able to infer the opportunity cost in each country of an additional year of expected life. Other important determinants of average longevity include environmental conditions, social customs and income distribution. Although our approach does not directly value these distributional and environmental aspects of living conditions, they influence our welfare comparisons inasmuch as they impinge on observed longevity. We find that international data on life expectancy and GDP per capita allow the construction of a partial ordering of the welfare of the 'representative agent' of each country.

In Section IV, we test whether either of the two most common rankings of national welfare, by real GDP per capita and by the Human Development Index (HDI), satisfy the partial ordering of our revealed preference analysis. Not surprisingly, we find that the ranking of real GDP per capita fails to satisfy the revealed preference criteria in a number of cases where levels of GDP per capita are fairly close, but life expectancy is substantially higher in the lower-ranked country (for example, Belgium and the Netherlands). Perhaps more surprisingly, we find that the HDI ranking, amended to provide a fair comparison, also fails to satisfy the criterion of revealed preference. The HDI reverses the 'mistakes' of the GDP rankings, but, by our criteria, it overcompensates. For example, the HDI method places the Netherlands not only above Belgium but also above the United States. This HDI ranking contradicts our finding that the United States could have afforded the average Dutch consumption bundle and, at the same time, could have channeled sufficient resources into health and education to surpass Dutch longevity, whereas the Dutch could not have achieved US outcomes.

Since neither GDP nor HDI rankings satisfy our revealed preference test, we present in Section V a method of constructing a utility-consistent index which does satisfy the test, based on the true Afriat index developed by Dowrick and Quiggin (1997). Finally, in Section VI, we consider possible extensions and modifications of the analysis.

I. Extended measures of GDP and multiple indicators

The limitations of GDP measures as a basis for international and intertemporal comparisons of living standards have been discussed on many occasions. One response has been to extend and adjust the standard GDP measure in various ways, such as by constructing measures in which corrective or defensive expenditures are excluded and household activities are included. Eisner (1988) has reviewed such extended systems of national accounts.

An alternative approach, particularly popular in the development literature, has been to report multiple indicators of social development including life expectancy, literacy rates and, occasionally, measures of personal freedom and political democracy. Dasgupta (1988) has argued cogently that indicators such as life expectancy and literacy rates need to be considered as important indicators of development within a desire-fulfillment framework.

Multiple indicators have sometimes been aggregated into alternative indicators such as the United Nations Development Programme’s (1990, 1994, 1995) Human Development Index (HDI), based on life expectancy, literacy and GDP per capita (Desai 1991). In most cases, the aggregation is based on a scaling procedure designed to reflect human needs and a declining marginal utility of consumption (See, for example, Anand and Sen 1994.) and an *a priori* assignment of weights.

In some cases, alternative aggregation procedures have been based on the statistical properties of the data. For example, Slottje et al. (1991) form indexes using principal components and other measures of statistical association within their data sets, and Noorbakhsh (1998) uses principal component analysis to modify the HDI.

Whilst we may lack confidence in the manner in which these alternative measures extend the notion of GDP, there is considerable merit in the objective of deriving measures which more closely approximate a broad concept of welfare. Such measures should improve attempts to evaluate policies where benefits are diffused widely. Barro and Lee (1993, 1994) attempt to measure the social value of investment in human capital through schooling. Others have tried to quantify the social returns to public investment in the core infrastructure of roads, water and sewerage (for example Aschauer 1989; Easterly and Rebelo 1993.).

These studies are based on the premise that there may be diffuse but significant spillover benefits from particular investments, benefits which will not be picked up in analysis of private returns or of the productivity of individual firms. These studies test for external effects in broad aggregates such as GDP and try to quantify social rates of return. Implicit in these tests is the notion that the GDP aggregates are inclusive and exhaustive of those benefits. But the very nature of the potential spillovers in these examples highlight the inadequacies of aggregate measures which ignore health and environmental outcomes. In particular, it is important to consider the direct effects of education in improving health and, ultimately, life expectancy (Kenkel 1991).

The weighting procedures used to derive the HDI are based on judgement rather than on welfare theory. For Desai (1991, pp. 353-4) this does not constitute a problem since:

> *Social reciprocity, participation, interdependence are put into the wellbeing function. It is perhaps possible though hardly worthwhile to translate back many of the things listed above into commodity terms since the externality/interdependence valuation so difficult in consumer theory remains. Thus the concept of human development deliberately goes beyond the utilitarian calculus.*

This argument carries weight at the level of individual welfare measurement. However, composite indicators such as the HDI have, in practice, been developed and publicized at a national aggregate level. At the aggregate level of national GDP, we suggest that it is feasible to use welfare theory to measure the trade-offs between indicators of social capability and consumption of commodities, and thereby to estimate the weights implicit in current social choices. So it may be worthwhile to attempt the ‘utilitarian calculus’, if only to determine how the relative weights of the calculus relate to the explicit, but essentially arbitrary, weights used in current composite measures.

We suggest that our multilateral index number method, based on economic welfare theory (‘the utilitarian calculus’), is an interesting and illuminating alternative to the essentially *ad hoc* weighting and scaling procedures of the HDI. The methods we demonstrate for combining measures of GDP and life expectancy can be extended to include other important social indicators such as literacy and inequality.

II. Revealed Preference Analysis

Standard revealed preference analysis is based on the assumption that for each country i, we have a vector of relative prices $\mathbf{P}^i$, where the numéraire is the international average consumption bundle, and a vector of quantities $\mathbf{Q}^i$ representing per capita consumption of each group of goods. It is assumed that, for the representative individual, the budget set is given by:

(1) $$\{\mathbf{Q}: \mathbf{P}^i \cdot \mathbf{Q} \leq \mathbf{P}^i \cdot \mathbf{Q}^i\},$$

and that the consumption vector $\mathbf{Q}^i$ is the most preferred element of this set.

The usual revealed preference method cannot be applied to social indicators such as health, life expectancy and literacy rates because of the absence of a natural price measure. However, revealed preference principles may still be applied if the indicators are systematically related to the consumption of market goods: for example, if health outcomes can be regarded as outputs with consumption of food and medical services as inputs.

We assume that the vector of market goods may be partitioned into two sub-vectors $\mathbf{Q}^1$, $\mathbf{Q}^2$ where the elements of $\mathbf{Q}^1$ are goods valued in their own right, and the elements of $\mathbf{Q}^2$ are valued only as inputs to the production of the final goods that are captured by the social indicators. Some elements of $\mathbf{Q}^1$ may also contribute to the production of the vector of social indicators which we denote $\mathbf{Z}$. It is assumed that for each country i, there exists a technology represented by a set T^i:

(2) $$T^i = \{(\mathbf{Q}^1, \mathbf{Q}^2, \mathbf{Z}) : \mathbf{Z} \text{ is attainable given the inputs } \mathbf{Q}^1, \mathbf{Q}^2\}.$$

The feasible consumption set for country i is given by:

(3) $$X^i = \{(\mathbf{Q}^1, \mathbf{Q}^2, \mathbf{Z}) \quad T^i: \mathbf{P}^i \cdot \mathbf{Q} \leq \mathbf{P}^i \cdot \mathbf{Q}^i, \text{ where } \mathbf{Q} = (\mathbf{Q}^1, \mathbf{Q}^2)\}.$$

Now consider two countries, A and B, with price–quantity vectors given by $(\mathbf{Q}^{1A}, \mathbf{Q}^{2A}, \mathbf{Z}^{A}, \mathbf{P}^{A})$, $(\mathbf{Q}^{1B}, \mathbf{Q}^{2B}, \mathbf{Z}^{B}, \mathbf{P}^{B})$. Country A is revealed preferred to Country B if:

(i) $\mathbf{Q}^{2}, (\mathbf{Q}^{1B}, \mathbf{Q}^{2}, \mathbf{Z}^{B}) \quad X^{A}$,

that is to say, if A could have afforded both B's bundle of goods and the inputs required to match B's social indicators.

Similarly, B is revealed preferred to A if:

(ii) $\mathbf{Q}^{2}, (\mathbf{Q}^{1A}, \mathbf{Q}^{2}, \mathbf{Z}^{A}) \quad X^{B}$.

If neither (i) nor (ii) holds, no ranking is possible in the absence of additional information on preferences. If both (i) and (ii) hold, one of the maintained hypotheses, such as the hypothesis of common tastes, must be violated.

The technology T^{1} defines the maximum indicator function, where L is one of the social indicators, an element of $\mathbf{Z}$, and $\mathbf{Z}^{-}$ is the vector of other indicators:

$$L^{A,B}\ (\mathbf{Q}^{1B},\ \mathbf{Z}^{B},\ \mathbf{P}^{A}) \ = \ \max\{L:\ \mathbf{Q} \geq (\mathbf{Q}^{1B},\ \mathbf{0}),\ \mathbf{Z}^{-} \geq \mathbf{Z}^{-B},\ (\mathbf{Q},\ \mathbf{Z}) \quad \mathbf{X}_{A}\}.$$

In the example we apply in the next section, $L^{A,B}$ is the maximum attainable life expectancy for Country A, subject to the constraint that Country A achieves at least the same consumption of goods and other social indicators as Country B. The condition that Country A is revealed preferred to Country B is then:

(4) $$L^{A,B} \ \geq \ L^{B}.$$

This welfare-theoretic approach to a comparison of international consumption levels and life expectancy poses both philosophical and practical problems. The notion that tastes are common across the world appears fanciful when we know that our families and colleagues display widely-differing preferences. Furthermore, the practice of comparing average national outcomes may seem distasteful or irrelevant to those whose prime interest lies in individual variety, particularly those concerned with inequality in social and economic outcomes. Nevertheless, given the frequent use of international comparisons of measures of average income, life expectancy and so on, it is desirable that such comparisons should have a clear economic interpretation. National and international expenditure patterns happen to be consistent with the hypothesis of rational choice by a representative agent. The use of this information to make international comparisons of living standards does not imply that individual preferences can be aggregated into those of a representative national consumer, nor that tastes are common worldwide. What it does imply is that such comparisons can be undertaken and interpreted using standard economic principles.

The principal practical problem in applying revealed preference tests to social indicators is that we do not observe market valuations of these indicators. Varian (1988) has shown that the absence of even one price renders the standard revealed preference approach inapplicable. We argue that this problem may be overcome if it is possible to estimate a household production technology in which outcomes relevant to utility are determined, in part, by consumption of market goods such as food and health services. Given knowledge of such a technology and the relevant input prices, we demonstrate that it is possible to apply the theory of revealed preference to the resulting feasible consumption sets. Our procedure takes account of the possibility that some of the inputs such as food may be valued in their own right whilst other inputs such as medical expenditures may be purely instrumental. We also take account of the possibility that the implicit price of an indicator such as life expectancy is unlikely to be linear.

III. Cross country comparisons based on disaggregated GDP and life expectancy

Our primary sources of GDP data are the two parts of the report on Phase IV of the International Comparison Project (United Nations and Commission of the European Community 1986,1987). Part One presents the aggregated ‘league tables’ of GDP per capita, evaluated at international prices for 60 countries in 1980. Part Two contains the details of prices and quantities disaggregated into private consumption, government consumption and capital formation. These broad expenditure categories are broken down into twelve categories which are further disaggregated into 38 more detailed categories. Life expectancy is the average for the period 1975–80, taken from United Nations (1989). This measure is based on the notional expected life of newborn infants if they were to face current age-specific mortality rates throughout their future lives. It is the single most widely-used measure of the health of a nation. The absence of demographic data for Botswana reduces the sample to 59 countries. We have also excluded Guatemala from our sample because we do not believe the data on medical expenditures can be correct; it suggests that Guatemala spent over 20 per cent of GDP on health care, some three or four times as much as other countries at a similar stage of development. This reduces our sample size to 58. The data set is summarized in Table 1. GDP is the usual aggregate converted into $US using purchasing power parity (PPP) exchange rates. [2]

Figure 1, which is a scatter plot of life expectancy against GDP per capita, shows that countries with higher GDP per capita tend to have higher life expectancy. The relationship is strongly non-linear. Whereas the rank correlation is 0. 93, the linear correlation coefficient is only 0. 82. The figure suggests a semi-logarithmic relationship; when GDP per capita is scaled logarithmically, the correlation coefficient with life expectancy rises to 0. 93.

The relationship between GDP per capita and life expectancy is strong, but there are many interesting variations. The United States, for example, is the second richest country in the sample but its life expectancy of 73. 2 years is one year below that of Canada and two years below that of Japan, Norway and the Netherlands. Luxembourg and Germany also rank significantly lower in life expectancy than in GDP per capita . On the other hand, Sri Lanka, Spain and Greece stand out as countries which have much longer life expectancy than their GDP ranking might suggest. For these three long-lived populations, life expectancy is comparable to that typically found in countries which have twice the level of GDP per capita.

These disparities highlight the nature of the problem of constructing a single measure or ranking of living standards. On standard comparisons of GDP per capita, the United States appears to be much better off than Japan, and Brazil appears to be much better off than Sri Lanka. These rankings will be reversed if sufficient weight is attached to life expectancy. The choice of these weights is usually quite subjective, as is the choice of whether to enter GDP per capita in linear or logarithmic form.

Our revealed preference approach requires knowledge of the production technology for

---

[2] The detailed data on prices and quantities of GDP components is available on request from the authors and also available by electronic access.

our single social indicator, life expectancy, denoted L. A number of approaches to the estimation of such a technology might be considered. First, there is a large and growing literature on the cost-effectiveness of particular medical interventions, most commonly assessed in terms of the marginal cost of gaining quality-adjusted life-year (QALY). Applying the results of this literature to health interventions which are marginal in a given country yields estimates of the marginal cost of increasing life expectancy. (Note that in this case we would ideally want data on internationally comparable measures of quality-adjusted life-expectancy.)

Second, data on the compensating wage differentials associated with increases in the risk of work-related illness, death and injury can be used to estimate the social costs of life-saving interventions. The assumption underlying this approach is that employers should willing to pay at least as much for a given improvement in occupational health as they would otherwise pay in compensating differentials to workers for the associated health risk.

In this paper, a third approach is adopted, based on econometric estimates of the production technology for health, derived from the International Comparisons Project data set (Dowrick, Dunlop and Quiggin 1998). Since there is only a single social indicator, the technology may be represented by a production function:

$$L = f(\mathbf{Q}^1, H), \tag{5}$$

where Q is the vector of consumption goods other than health, and H is consumption of health services, assumed to be valued only instrumentally. All variables are expressed in per capita terms.

Dowrick, Dunlop and Quiggin (1998) estimate[3] the production function for life expectancy as:

$$L^i = 7.45 + 2.55\log(H^i) + 3.47\log(E^i) + 3.75\log(F^i) + \varepsilon_i. , \tag{6}$$

[3] Dowrick, Dunlop and Quiggin (1998, Table 3) report two specifications, one of which allows for a smaller impact of food expenditure in more developed countries. In this paper, we focus on the simpler loglinear specification.

where E is consumption of education services and F is consumption of food. More detailed results are presented in Table 2.

The country-specific component of the production technology is captured by the disturbance term ε_i. A positive value of ε_i indicates that for given levels of consumption of health services, education services and food, country i has higher-than-average life expectancy. These residuals might capture differences in climate, public health infrastructure, and income distribution. The residuals, along with data on L, H, E and F, and the relative prices of the inputs, are listed in Appendix Table 1.

The main interest of Dowrick, Dunlop and Quiggin (1998) was the derivation of internationally comparable estimates of the cost of raising life expectancy or equivalently, of the cost of saving the life of an individual of a specified age. Using equation (6), the cost of saving the life of a newborn individual in a wealthy country was estimated at between $US1 million and $US2 million (at 1980 prices), which is comparable to estimates of the subjective value of life derived from studies of wage premiums (Hwang, Reed and Hubbard 1992) and from applications of the contingent valuation method (Viscusi 1990).

Thus, for developed countries, the results of the cross-country econometric approach agree with those of intra-country approaches. The main advantage of the cross-country econometric approach, for our purposes, is that it provides consistent estimates for a large number of countries of the opportunity cost of devoting resources to extending life expectancy.

Given the estimated relationship (6), it is a straightforward exercise to determine for any given pair of countries, A and B, the maximum life expectancy attainable by country A subject to the constraint that consumption of all goods other than health services is greater than or equal to that of country B. Subject to a budget constraint, we choose H, E and F to maximize A's predicted life expectancy as given by (6).

The maximization problem may be written as:

(7) $L^{AB} = \text{Max } L^{A}(H, E, F)$, subject to the constraints

i) $E \geq E^{B}$;

ii) $F \geq F^B$;

iii) $H \geq 0$; and

iv) $P^A_E(E - E^B) + P^A_F(F - F^B) + P^A_H(H - H^B) \leq \mathbf{P}^A \cdot \mathbf{Q}^A - \mathbf{P}^A \cdot \mathbf{Q}^B \equiv S^{AB}$.

The final constraint is simply that country A can spend more than country B on education, health and food only if there is a surplus, denoted S^{AB}, after consuming country B's goods. Given that the production function is strictly increasing in all its arguments, the budget constraint will always be binding. Further, the semi-log form ensures that health expenditures must always be strictly positive. Thus, constraint iii) will never bind, and constraint iv) will always bind.

It is possible, however, that the constraints on education and food may be either binding or slack. Predicted life expectancy might be increased by transferring resources from, say, food to health, but we cannot carry out a revealed preference comparison if food consumption falls below that in the reference country. Accordingly we can limit the feasible solutions to (8) to the following four cases:

a) constraints i) and ii) bind;

b) constraint i) is binding, constraint ii) is slack;

c) constraint ii) is binding, constraint i) is slack; and

d) constraints i) and ii) are slack.

Given the semi-log form of the budget constraint, it comes as no surprise that the solution under each of these cases can be expressed as a simple sharing rule. First calculate ‘disposable income’, the maximum expenditure available after matching the other country's consumption of goods for which the constraints are binding. Then allocate this disposable income amongst health, education and food (or the subset of these for which constraints are not binding). The budget shares for this allocation are in direct proportion to the regression coefficients, α_H, α_E, α_F in the production function.

In case d) where neither E nor F is constrained, the optimal allocation of resources to health, for example, is given by:

(8) $$H^{d,AB} = (S^{AB} + P^A_H H^B + P^A_F F^B + P^A_E E^B) \cdot (\alpha_H / [\alpha_H + \alpha_F + \alpha_E]) / P^A_H.$$

The sharing rule for E^d and F^d is identical in form. Substitution of these optimal expenditures into (7) yields $L^{d,AB}$. A similar sharing rule applies to the division of disposable income between health and education alone, as in case c), or between health and food, as in case b). In case a), all disposable income is spent on health. It is then straightforward to compute the maximum feasible life expectancy $L^{AB} = \max\{L^{a,AB}, L^{b,AB}, L^{c,AB}, L^{d,AB}\}$. If this life expectancy is greater than or equal to that of country B, A is revealed preferred to B.

IV. Revealed Preference Results

Results of the revealed preference exercise are summarized in the matrix of Table 2 where element ij = '+' indicates that i is revealed preferred to j, '-' indicates the reverse, and 'nc' indicates that the revealed preference test is inconclusive. There is one case where the Generalized Axiom of Revealed Preference discussed by Varian (1982, 1983), enables us to resolve a comparison which is indeterminate using the Weak Axiom of Revealed Preference : Luxembourg and Japan are not directly comparable on the basis of the Weak Axiom, but Luxembourg is revealed preferred to France, and France is revealed preferred to Japan. In this case the symbols G+ and G- in Table 2 indicate the rankings derived from the Generalized Axiom. There is only one instance where we are forced to reject the maintained hypothesis of common tastes, using the Weak Axiom, that is in the comparison between Finland and Austria. This is represented as '!!' in Table 2. There are no further rejections of common tastes under the Generalized Axiom.

The maintained hypothesis of common tastes for the set of countries excluding Finland may be checked using the Generalized Axiom, following Lemma 1 of Dowrick and Quiggin (1994). To perform this test, the rows of Table 2 may be ordered according to the ICP rankings, adjusted for unambiguous reversals. It is then simply necessary to check that there are no plus signs below the diagonal. This test reveals no violations of the Generalized Axiom, with the exception of the previously observed violations of Weak Axiom. These results may be compared with those of Dowrick and Quiggin (1994) who used the ICP consumption data without life expectancy, treating medical services as a final good. A notable difference is an increase in the number of noncomparable pairs, marked 'nc' in Table 5, from 77 to 88.

Countries are ranked in Table 2 by GDP per capita, using the ICP measure. There are six pairwise comparisons where the ICP ranking contradicts our revealed preference partial ordering. These cases are marked ‘(+)’. For instance, The Netherlands (GDP per capita = $9320) is revealed preferred to Belgium (GDP per capita = $9435). Japan, the United Kingdom and Italy are each revealed preferred to Finland, even though GDP per capita in Finland is higher; similarly, Greece is revealed preferred to Venezuela, and Zimbabwe is revealed preferred to Nigeria.

We can go further than these outright rejections of GDP rankings to consider the ‘ties’, those cases where countries cannot be separated by revealed preference. The ICP rankings understate the relative standard of living in countries such as Norway, Belgium, Italy, Sri Lanka and Tanzania which rank higher on life expectancy than on GDP per capita. Countries which are several places higher on the ICP rankings and are revealed preferred on the standard GDP data are not revealed preferred when life expectancy is included in the comparison. Most notably, Peru is not revealed preferred to Sri Lanka, although, on the ICP measure, Peru is more than twice as well off. Neither is Bolivia, with an ICP measure of $1633, revealed preferred to Zimbabwe with an ICP measure of only $895.

Equally, the standard of living appears to be overstated by GDP measures which ignore the comparatively low life expectancy of both Peru and Bolivia and other countries such as Luxembourg, Belgium, Brazil, Senegal and Malawi. All of these countries are non-revealed preferred to countries which are lower on the ICP rankings.

Does the Human Development Index, which incorporates life expectancy and GDP per capita in a composite index, yield rankings which are more consistent with our revealed preference ordering? In order to provide a fair test we follow the method of constructing the index described in the Human Development Report (United Nations Development Program 1990, p. 109), but use only GDP per capita and life expectancy, not literacy. [4] This gives a

---

[4] Later definitions of the HDI, such as that presented in United Nations Development Program (1994, p. 108), truncate the upper end of the GDP range and add other variables such as years of schooling. We choose the simple definition which is closest to the spirit of our revealed preference test.

measure:

(9) HDI = [log(GDP) - log(284)]/2 log(11619/284) + [LE-40]/2 [75. 5-40]

where 284 and 11619 are the minimum and maximum observed levels of GDP per capita and 40 and 75. 5 are the limits observed on life expectancy.

This measure of the HDI is listed in Table 4 and used as the ranking for the revealed preference matrix. The six mis-rankings found in the comparisons based on GDP are corrected by the inclusion of life expectancy in the HDI . For example, the Netherlands is now ranked above Belgium, in line with the revealed preference test. But the HDI ranking is biased in the opposite direction: the Netherlands is ranked above the United States, Denmark Germany and Luxembourg although those four countries are each revealed preferred. There are 27 pairwise comparisons where the HDI rankings contradict revealed preference.

Since neither GDP nor HDI rankings satisfy our revealed preference test, the challenge of the next section is to provide a method of constructing a utility-consistent index which does satisfy the test.

V. Construction of a true index of GDP per capita incorporating life expectancy

In order to make cardinal comparisons we need to restrict the form of the notional preference relationship. One approach would be to estimate the parameters of a specific functional form for a utility or expenditure function, an approach surveyed by Diewert and Nakamura (1993). The alternative, which we prefer, is to require that the preference relationship be homothetic. This is a necessary and sufficient condition to yield a non-trivial money-metric utility independent of any particular reference price vector. Whilst homotheticity may appear to be a strong condition, it has been found that both time series data for US GDP and cross-country GDP data can be so represented (Manser and McDonald (1988; Dowrick and Quiggin 1997).

Dowrick & Quiggin (1997) have demonstrated the construction and multilateral properties of true indexes, defined as money-metric utility numbers where each pairwise comparison lies within the Paasche and Laspeyres bounds. In order to incorporate social indicators into this index number framework, however, we need a linear approximation to the price of the indicator. We achieve this by substituting into the standard GDP bundle that notional quantity of medical services which would be predicted to have delivered the observed longevity. This procedure has the effect of transforming the units of life expectancy, for which the cost is highly non-linear, into units of equivalent medical services with a constant price.

Use of this approximation allows us to test the GDP and life expectancy data for homotheticity and derive a set of ideal true index numbers for the 53 countries which fit a common homothetic representation. We also derive generalized Fisher index numbers for those five countries outside the homothetic set and find that these index numbers do fit the revealed preference ordering. Since they are based on a homothetic representation, they have a meaningful cardinal interpretation: If A's index number is twice that of B's, we infer that the representative household would be indifferent between, on the one hand, the actual GDP bundle and life expectancy of A, and, on the other hand, twice the GDP bundle of B with twice the notional health expenditures of B.

Our first step involves estimating the notional quantity of medical services which would explain observed life expectancy. From (6) this is calculated for country i as H_i^* such that $f(H_i^*, E_i, F_i) = L_i$ where E, F, and L are observed values. Where the life expectancy residual in (6) is positive, notional medical services are greater than measured services. We treat the data as if the true quality of medical services in that country was underestimated, so where the notional quantity exceeds the measured, we reduce the measured price of medical services proportionally to yield a notional price $P^{H}_{i}{}^{*}$.

We can then follow the procedure described by Dowrick and Quiggin (1997). The first step is to calculate the matrix, **L**, of Laspeyres ratios for each pairwise comparison. If the generating preference relationship is homothetic, the true utility ratio, u_i/u_j, must lie below the Laspeyres ratio, L_{ij} . We can test whether there exists a set of such utility numbers, $u_1, u_2, \ldots, u_n$ that satisfies all the bilateral conditions. Following Varian's (1983) algorithm, we replace each element L_{ij} by its minimum path, M_{ij}, to construct the minimum path matrix, **M**, where :

$$M_{ij} = \min._{(k,\ldots,m)} [L_{ij}, (L_{ik} + L_{kl} + \ \ldots + L_{mj})] .$$

The element M_{ij} gives the upper homothetic bound for the logarithm of j's utility index relative to i's. Afriat (1981) shows that a homothetic representation is possible if and only if the diagonals of this matrix, M_{jj} , are non-negative.

Applying this technique to our data, we find negative diagonals for our full sample of 58 countries, implying that a homothetic representation is not possible for the full set. However, by dropping five countries - Finland, Israel, Hungary, Poland and Paraguay - we are able to achieve a homothetic representation for the remaining set of 53 countries.

The row and (negative) column averages of the minimum path matrix give, respectively, the lower and upper bounds to the true index of country i relative to the mean of the index. These bounds are listed in Table 5 in logarithmic form, so that a number of 0. 2, for example, indicates that the bound is $e^{0.2}$ =1. 22 times the sample mean. The sets of upper and lower bounds each constitute a true index. We also list the mid-point of the bounds which constitute the Ideal Afriat Index (a true index which reduces to the Fisher ideal index in the case of only two observations). This gives our preferred true index, **A=** $\{A_1, A_2, \ldots, A_h\}$, over the set of observations, **h**, for which a homothetic representation is possible.

Dealing with the five 'non-homothetic' countries is a problem if we want to establish a method that can be applied to all countries. One solution is to follow Varian (1983) in defining a utility function for any consumption bundle, $\mathbf{x}_k$:

$$P(\mathbf{x}_k) = \min_{i\ \ \mathbf{h}} (A_i + \ln(\mathbf{p}_i\mathbf{x}_k / \mathbf{p}_i\mathbf{x}_i)).$$

This utility function yields A_k for any observation within **h** and the minimum of the set of Paasche valuations (using i's prices to value $\mathbf{x}_k$ relative to each $\mathbf{x}_i$) for any observation k outside **h.** An obvious counterpart is to define the utility function:

$$L(\mathbf{x}_k) = \max_{i\ \ \mathbf{h}} (A_i + \ln(\mathbf{p}_k\mathbf{x}_k / \mathbf{p}_k\mathbf{x}_i)),$$

which yields A_i for any observation k within **h** and gives the maximum of the Laspeyres valuations (using country k's own prices) for k outside **h**. A natural extension is to define the mid-point of these two functions,

$$A^*(\mathbf{x}) = [P(\mathbf{x}_k) + L(\mathbf{x}_k)]/ 2,$$

which extends the true index A to include observations which do not fit the homothetic representation. The non-homothetic numbers A^*_k can still be given an economic interpretation. They give the utility of the representative agent, whose preferences are homothetic and consistent with observations in the set **h**, if they were allocated the consumption bundle $\mathbf{x}_k$ **-** even though they would not have chosen the bundle $\mathbf{x}_k$ at prices $\mathbf{p}_k$.

The index **A*** is listed in Table 4. The five countries with non-homothetic observations are marked by an asterisk. The countries are listed in descending order of A_i. We find that these rankings are fully consistent with the partial ordering of the revealed preference tests. For example, the Netherlands is now ranked above Belgium (unlike the GDP rankings) but below the United States and Denmark (unlike the HDI rankings).

For purposes of comparison, GDP per capita is also listed in Table 4 in logarithmic form, normalized to a zero mean for the sample of 53 countries. We find that in one-third of the cases the log GDP index lies outside the true bounds. For example, on the log index, the Netherlands is only 1. 30 above the mean, but the true bounds are (1. 33, 1. 48).

Figure 2 displays both GDP per capita and the index **A*,** measured on the left-hand axis, and the HDI measured on the right-hand axis. The countries are ordered by **A***, so non-monotonicity of the GDP or HDI line indicates a change in the ordering relative to **A***. Figure 2 illustrates the tendency for countries, like the Netherlands, which are ranked too low by GDP to be ranked too high by the HDI, and *vice versa*.

VI. **Possible Extensions and Modifications**

Although our empirical analysis has been undertaken for the case of a single social indicator, the general method set out above is applicable to a vector of indicators.

Apart from life-expectancy, the social indicators most frequently used are literacy rates and infant mortality. The inclusion of literacy rates as a function of consumption of education services and possibly other goods is a straightforward extension. The inclusion of infant mortality as a measure in addition to life expectancy seems dubious since life expectancy is itself determined by age-specific death rates. It would seem preferable to disaggregate by measuring age-specific death rates for a number of age-groups.

Another possible extension would be to replace life expectancy with a measure such as quality-adjusted-life-years (QALYs), incorporating multiple measures of health outcomes. This would take account of morbidity as well as mortality and thereby capture a wider range of outputs produced using health services.

An alternative, more in keeping with the revealed preference approach, would be to use the multiple measures of health outcomes employed to form the QALY index as elements of the vector **Z**. Data on health outcomes and consumption would form the basis for estimation of a cost function.

One advantage of using indexes that include social indicators is that they give more weight to ‘basic goods’ and hence to the well-being of the poor. This might be seen as a reason for preferring measures such as life expectancy and infant mortality, which are based on average age-specific death rates, to broader measures of average health status. But the approach most compatible with revealed preference theory would be to analyse household data and disaggregate by household consumption levels. It would then be possible to examine the actual living standards of the poor rather than relying on potentially misleading measures of income inequality. Such an approach would lend itself naturally to the incorporation of indicators such as mortality and morbidity rates.

The main limitation of the revealed preference approach is that it relies on the assumption that welfare is based on the consumption of goods and services. It may be argued that social indicators are categorically different from the consumption items incorporated in GDP, so that no coherent index that aggregated them could be formed. We argue that it is not necessary to accept this argument in relation to literacy or life-expectancy. Existing GDP indexes incorporate health and education services, so the only relevant distinction is between inputs and outputs. There is a stronger argument to suggest that measures of political freedom are categorically distinct from measures of consumption.

VII. **Conclusions**

Two different types of objection have been made to the use of measures of GDP per capita as a basis for international comparisons of living standards. First, the set of consumption items that make up the GDP index omits important elements of wellbeing, such as health status. Second, standard procedures for combining quantities of individual goods into a GDP index at national or international prices do not typically reflect social opportunity costs, so that rankings derived from such indexes may be inconsistent with rankings inferred from revealed preferences.

In this paper, it has been shown that these difficulties may be overcome using a generalization of the revealed preference approach pioneered by Samuelson (1947). This generalization provides a natural framework for the incorporation of information on social indicators into revealed preference analysis. We find that the international data on life expectancy and GDP per capita allow the construction of a partial ordering of the welfare of a 'representative agent' facing the prices and outcome of each country.

We have then constructed a true index (with well-defined upper and lower bounds) which incorporates both the standard components of GDP and life expectancy. Although this method requires that we redefine the utility function on 'notional health services' rather than directly on life expectancy, the resulting ranking is fully consistent with the partial ordering of revealed preference. Where a unique set of cardinal index numbers is required, this method is preferable to the relatively arbitrary weighting procedures of the Human Development Index and other composite indices.

These new index numbers incorporate life expectancy into measures of national income and welfare for the purposes of quantitative policy analysis. The methods which we have developed, both for revealed preference rankings and for the generation of cardinal index numbers, can, in principle, be extended to include other social indicators such as literacy. It may be possible to also take account of inequality in income and inequality in life expectancy using explicit social welfare functions. However, as argued by Desai (1991), there remains a class of indicators of human welfare, including notions of freedom and democracy and community, which are probably not amenable to the 'utilitarian calculus'.

Our index number approach should allow improved testing of models of economic development in which economic welfare is are defined more widely than by conventional GDP. For example, our cross-country analysis confirms the findings of intra-country studies that education yields both direct and indirect benefits for health in general and life expectancy in particular. This finding suggests that estimates of the impact of human capital investment on economic growth may understate the welfare benefits if they measure growth solely in terms of GDP and ignore important indicators such as life expectancy. Equally, this method may have important implications in testing the relationship between measures of inequality and aggregate economic development, since reductions in inequality are likely to increase longevity. Application of our method to a panel of observations on GDP, life expectancy and inequality would enable these hypotheses to be tested.

References

Afriat, S. N. (1981) ‘On the constructability of consistent price indices between several periods simultaneously’, in A. Deaton, ed. , *Essays in the Theory and Measurement of Consumer Behaviour*, 133–161, Cambridge: Cambridge University Press.

Anand, Sudhir and Amartya K. Sen (1994), ‘Human development index: methodology and measurement’, *Occasional Papers* No. 12, New York: UNDP Human Development Report Office.

Aschauer, D. (1988), ‘Is public expenditure productive?’, *Journal of Monetary Economics* 23(2), 177–200.

Barro, Robert J. and Jong Wha Lee (1993), ‘International comparisons of educational attainment’, *Journal of Monetary Economics* 32, 363–94.

Barro, Robert J. and Jong Wha Lee (1994). “Sources of economic growth. ” *Carnegie Rochester Conference Series on Public Policy* 40(0): 1–46.

Dasgupta, Ajit K. (1988), *Growth, Development and Welfare*, Oxford: Basil Blackwell.

Desai, Meghnad (1991), ‘Human Development: concepts and measurement’, *European Economic Review* 35, 350–357.

Diewert, W. Erwin and Alice O. Nakamura (1993), eds, *Essays in Index Number Theory*, Amsterdam: North–Holland.

Dowrick, Steve and John Quiggin (1994), ‘International comparisons of living standards and tastes: a revealed preference analysis’, *American Economic Review* 84(1), 332–41.

Dowrick, Steve and John Quiggin (1997), ‘True measures of GDP and convergence’, *American Economic Review* 87(1), 41–64.

Dowrick, Steve, Yvonne Dunlop and John Quiggin (1998), ‘The cost of life expectancy and the implicit social valuation of life’, *Scandinavian Journal of Economics* 100(4), 673–691.

Easterly, William and Sergio Rebelo (1993). "Fiscal policy and economic growth: an empirical investigation." *Journal of Monetary Economics* 32(3): 417–458.

Eisner, Robert (1988), ‘Extended accounts for national income and product’, *Journal of Economic Literature* 26 (4): 1611–84.

Hwang, H. S. , W. R. Reed, and C. Hubbard, (1992), ‘Compensating wage differentials and unobserved productivity’, Journal of Political Economy 100, 835–58.

Kenkel, Donald S. (1991), ‘Health behaviour, health knowledge and schooling’, *Journal of Political Economy* 99 (2), 287–305.

Kravis, Irving, Alan Heston, and Robert Summers, (1982), *World Product and Income: International Comparisons of Gross Product*, Baltimore, MD: John Hopkins University Press.

Manser, Marilyn E. and Richard J. McDonald (1988), ‘An analysis of substitution bias in measuring inflation, 1959–85’, *Econometrica* 56(4), 909–930.

Noorbakhsh, Farhad (1998), ‘A modified human development index’, *World Development* 26(3), 517–28.

OECD (1992), *Real expenditures in OECD countries: purchasing power parities for 1990*, Paris: OECD.

Samuelson, P. (1947), *Foundations of Economic Analysis*, Cambridge, Massachusetts: Harvard University Press,.

Scott, Maurice (1990), "Extended accounts for national income and product: a comment", *Journal of Economic Literature* 28 (3): 1172–86.

Slottje, Daniel J., Gerald W. Scully, Joseph G. Hirschberg and Kathy J. Hayes (1991), *Measuring the Quality of Life Across Countries: a multidimensional analysis*, Boulder, Colorado: Westview Press.

United Nations & Commission of the European Communities (1987), *World Comparisons of Purchasing Power and Real Product for 1980: Phase IV of the International Comparison Project: Part Two: Detailed Results for 60 Countries*, ST/ESA/STAT/SER. F/42(Part II), New York: United Nations.

United Nations (1989), *World Population Prospects 1988*, New York: United Nations.

United Nations Development Programme (1990), *Human Development Report 1990*, New York: Oxford University Press.

United Nations Development Programme (1994), *Human Development Report 1994*, New York: Oxford University Press.

United Nations Development Programme (1995), *Human Development Report 1995*, New York: Oxford University Press.

Varian, Hal R. (1982), ‘The nonparametric approach to demand analysis,’ *Econometrica*, 50, 945–73.

Varian, Hal R. (1983), 'Non–parametric tests of consumer behaviour', *Review of Economic Studies* L, 99–110.

Varian,. Hal R. (1988), ‘Revealed preference with a subset of goods’, *Journal of Economic Theory* 46(1),179–85.

Viscusi, K. (1990), ‘The value of risks to life and health’. *Journal of Economic Literature* 31, 1912–46.