THE CHOICE OF INDEX NUMBER: Part I

VALUATION AND EVALUATION

SANJAY G. REDDY*AND BENJAMIN PLENER†

Version 1.2.9 March 28, 2006

Abstract

This paper defines and discusses the choices that are implicit in alternative approaches to constructing index numbers. The central theoretical result is the exogeneity theorem, which states that an index number system possesses certain generally desirable properties, associated with meaningfulness, relevance and consistency if and only if it admits of a welfare interpretation that is invariant to the data to which it is applied. An index number system that does not permit a fixed welfare interpretation of this kind will fail to possess these properties. Moreover, it will not be possible to undertake meaningful multilateral comparisons between all possible units of comparison to which the index number system may be applied without adopting a meta-evaluative criterion that expresses welfare judgments concerning the relation between distinct evaluative standards. Therefore, explicit or implicit normative judgments are inescapable in the construction and use of index numbers intended to permit meaningful comparisons over an unrestricted domain.

We identify a set of bounds on the index numbers that an index number system produces (the "generalized within bounds" requirement) which define necessary and sufficient conditions for them to be rationalizable (i.e. to possess a welfare interpretation). However, multilateral index number systems that are widely in use violate this and other requirements of the exogeneity theorem. As a result, the index numbers they produce do not admit of a welfare interpretation which is invariant to the data to which they are applied, and indeed may not admit of any welfare

*Dept. of Economics, Barnard College, School of International and Public Affairs, and Institute for Social and Economic Research and Policy, Columbia University; sr793@columbia.edu

†Yale University; benjamin.plener@yale.edu

interpretation at all. The exogeneity theorem establishes that the rationale for "exact" index number systems is much less compelling than widely believed.

A fixed welfare interpretation is also a necessary and sufficient condition for the path-independence of the Divisia index. The theorem therefore establishes a three-way equivalence between desirable properties of continuous index number systems, desirable properties of discrete index number systems and the existence of a fixed welfare interpretation.

1 Introduction

Many questions of economic interest require evaluation and comparison across temporal, spatial or counterfactual units.[1] For example, one might wish to determine whether the volume of current economic activity is greater in Germany or in France, whether the cost of living in the United States was greater in 1939 or 1945, whether labour productivity would improve if Japan implemented a particular policy reform, or where in the world the incidence of poverty is greatest. Answering questions of this kind requires ordinal comparisons across the units being compared. Often, it may also be desired or found necessary to make cardinal comparisons. Unfortunately, the data we wish to compare (for example, patterns of production or consumption and price structures) are multi-dimensional. As such, it is not self-evident how to characterize alternative sets of data as possessing "more" or "less" of a given property. A binary relation of comparison may be imposed on the data that summarizes judgments concerning which data possesses "more" and which "less" of a given property. However, there are many ways to form such an ordering. If it is deemed necessary to make cardinal as well as ordinal statements, then it is necessary to associate level sets defined by this binary relation with real numbers – a process that we may refer to as scalarization. Ultimately, the problem of cardinal comparison is a problem of scalarization.[2]

The absence of a unique complete ordering relation in higher-dimensional spaces and the absence of a unique way to scalarize level sets defined by such a relation pose ubiquitous problems.

[1] We would like to offer our thanks to Sudhir Anand, Andre Burgstaller, John Collins, Duncan Foley, Wulf Gaertner, Julia Harrington, Jeff Helzner, Prabhjot Kaur, Sergey Kivalov, Eiichi Miyagawa, Lalith Munasinghe, Howard Nye, Thomas Pogge, Rajiv Sethi, Anwar Shaikh, Peter Spiegler and Michael Thaddeus for their invaluable help and above all to Amartya Sen, for showing us the way. We would also like to thank members of the economic theory seminar at Columbia University for their helpful comments.

[2] Scalarization understood in this sense is a broader idea than representability of a binary evaluative relation by a continuous real-valued function, which requires, as is well known, that particular technical conditions obtain.

Mathematically, we are faced with a problem of indeterminacy: there is more than one way to order vectors, and to scalarize level sets. Conceptually, we are faced with a problem of choice: which method of ordering and scalarization should we employ, and on what grounds should we justify this choice? We would like the ordinal and cardinal valuations that we undertake to have appropriate justification.[3]

Describing the properties of alternative methods of scalarization (and, *a fortiori*, ordering) and providing compelling justification for choices among them are the chief tasks of index number theory. An index number is an aggregate measure of some phenomenon. Two important examples are price and quantity indices, such as those which respectively convey the relative cost of living or the relative volume of economic activity between units of comparison. These units of comparison may be, for example, multiple countries at a point in time, a country at multiple points in time, or units in space-time such as multiple countries in multiple years.

2 Properties of Index Number Systems

In this section, we present a conceptual framework for the study of index number systems (methods of generating index numbers) and identify some potentially desirable properties which they may possess. Two of the properties are standard ones; transitivity and anonymity. Two of the properties relate to the welfare conception of index numbers; normative rationalizability and exogeneity. The final property that we study, bilateral comparison focus, is a requirement concerning the information that is permitted to influence the index number, and is a precise statement of the demand for "characteristicity" which is often made in the literature on index numbers. These properties are inexorably related in a manner which we shall make clear in the exogeneity theorem of section (4).

[3] It is interesting in this regard to note the perspective of Franklin M. Fisher and Karl Shell [*Economic Analysis of Production Price Indexes*, Cambridge: Cambridge University Press, 1998, p. xi]: "A price index, or index number, is a scalar representation of a list of prices relative to some base value. Why would one think of representing a vector by a scalar? Isn't this always a losing assignment? Why not work solely in terms of the most disaggregated model? There are several answers to these questions. First, there is no "most disaggregated model." For every breakdown of apples into their types, sizes, qualities etc. there is always a finer breakdown. Second, at highly disaggregated levels, price and quantity measurements are likely to be of lower quality. Hence, one might choose to work with more aggregated data while being mindful of the relationship between these data and less aggregated data. Third, the most appropriate economic model might be at a reasonably high level of aggregation. To use such a model requires the use of aggregated data, but it also requires an understanding of the relationship of these data to less aggregated data".

Index number theory has sometime been thought to have had three stages of development since its inception.[4] It has been proposed that early theorists thought of index numbers as statistical aggregates or measures of central tendency of a distribution. According to this "statistical approach" to index number theory, one index number should be chosen over another if it is a "better" descriptive statistic for the underlying distribution according to some criteria relevant to the choice of such a statistic. Irving Fisher pioneered what has been referred to in the recent literature as the "axiomatic approach" to index number theory, in which a method of generating index numbers is required to satisfy certain tests, or axiomatic properties. In this perspective, one index number should be selected over another if it satisfies a greater number of relevant tests (appropriately weighted by the analyst according to their importance).

More recently, theorists have shifted their focus from whether a given method of generating index numbers satisfies a series of such tests to whether the index numbers which it generates can be interpreted in terms of a welfare conception. In what has been called (perhaps overly ambitiously) the "economic approach" to index number theory a specific value of an index number is required to correspond to a level set in commodity space (in the case of a quantity index) or to the cost of achieving such a level set (in the case of a price index) , and this level set is meant to represent the preferences of a representative agent. It is interesting to note that a utility function is itself an instrument of scalarization, which maps the multidimensional space of commodity bundles to the real number line. Therefore, this literature has been concerned with identifying the relation between one mode of scalarizing commodity bundles and another. The "exactness" results of Erwin Diewert[5] and his predecessors[6] and the rationalizability theorem of Sydney Afriat[7] have provided the theoretical justification for a wide range of index numbers which are currently used. It is fair to say that the "economic approach" represents the presently most influential and indeed dominant approach in index number theory and practice. This paper is a critique and a generalization of the

[4] See for example Diewert, W.E. (1993a), "The Early History of Price Index Research", pp. 33-65 in *Essays in Index Number Theory*, Volume 1, W. E. Diewert and A. O. Nakamura (eds.), North-Holland, Amsterdam.

[5] See e.g. Diewert, W.E. (1976), "Exact and Superlative Index Numbers," *Journal of Econometrics* 4, 115-146; reprinted as pp.223-252 in *Essays in Index Number Theory, Volume 1*, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland, 1993.

[6] e.g. Konűs, A.A. (1925), translated and published as "The Problem of the True Index of the Cost of Living", *Econometrica* 7 (January) 1935, pp. 10-29, and Buyshgens (1925), "Sur une classe des hypersurfaces: A propos de 'l'index ideel' de M. Irving Fisher", *Mathematischkii Sbornik*, 32: 625-631.

[7] See e.g. Afriat (1977), *The Price Index* , Cambridge: Cambridge Univ. Press and the discussion in Varian (1983), "Non-parametric Tests of Consumer Behaviour", *Review of Economic Studies*, Vol. L, pp. 99-110.

so-called "economic approach" to index number theory.

3 Overview of the Problem and the Result

An individual index number defines a relation between two units of comparison (such as two points in space or in time). In particular, an index number identifies a factor of proportionality. This factor of proportionality must have some interpretation. For example, it can refer to the relative quantity of an outcome that is realized in different contexts, or it can refer to the relative quantity of a resource that is required in order to achieve the same outcome in different contexts. Index numbers must be defined in relation to some outcome if they are to have a meaningful interpretation. Attempts to choose the 'best' index number without reference to an outcome are fundamentally incoherent. The choice of the outcome in relation to which to understand an index number cannot itself be made on formal grounds. Rather, justifying this choice requires an evaluative judgement as to what outcome is of concern.[8]. In practice, the outcome with which index number theory has preeminently been concerned has been welfare achievement. Moreover, the concept of welfare achievement that has been adopted has been, even more narrowly, understood to refer to subjective preference satisfaction.

[8] "What index numbers are 'best'? Naturally much depends on the purpose in view." [Irving Fisher as quoted in "The Consumer Price Index and Index Number Purpose" (1999) by W. Erwin Diewert, available on http://www.unece.org/stats/documents/ces/ac.49/1999/crp.1.e.pdf]

Irving Fisher was far from clear or consistent in his statement of this view. He wrote seemingly contrarily in his "Prefatory Note" to *The Making of Index Numbers* (Boston: Houghton Mifflin, 1927), "To determine the pressure of steam, we do not take a popular vote: we consult a guage. Concerning a patient's temperature, we do not ask for opinions: we read a thermometer. In economics, however, as in education, though the need for measurement is as great as in physics or in medicine, we have been guided in the past largely by opinions. In the future we must substitute measurement. Toward this end, we must agree upon instruments of measurement. That is the subject of this book." Elsewhere (p.2) he writes that "We shall find that some of the formulae in general use and unhesitatingly accepted by uncritical users are really very inaccurate, while others have an extraordinary degree of precision", seemingly suggesting that precision is a concept that can be defined without regard to the purpose at hand.

The point that the appropriate index number depends on the purpose was recognized even by Lionel Robbins (notwithstanding his wish to purify economics of value judgments). Robbins states in his *Essay on the Nature and Significance of Economic Science* [Second Edition, Revised and Enlarged, 1935, London: MacMillan and Co., Ltd.]: "... there is no intention of denying the practical utility and significance of comparisons of certain prices over time, or the value of 'corrections' of these prices by suitably devised index numbers. It is not open to serious question that for certain questions of applied Economics on the one hand, and interpretation of history on the other, the index number technique is of great practical utility. Given a willingness to make arbitrary assumptions with regard to the significance of certain price sums, it is not denied that conclusions which are important for practice may be reached. All that is desired to emphasise is that such conclusions do not follow from the categories of pure theory, and that they must necessarily involve a conventional element depending either upon the assumption of a certain empirical constancy of data or upon arbitrary judgments of value with regard to the relative importance of particular prices and particular economic subjects."

It is reasonable to require that index numbers refer to some outcome as this is required if they are to possess a meaningful interpretation. We may refer to this reference outcome as an evaluative standard. If a set of index numbers corresponds to an evaluative standard, then it may be said that they are rationalized by the evaluative standard, or simply that they are rationalizable. If an index number system (a function used to generate index numbers from underlying data) always generates rationalizable index numbers, regardless of the data to which it is applied, then it may itself be called rationalizable.

If a finite set of index numbers is rationalizable, then there is a large (indeed infinite) number of evaluative standards that rationalizes it. If there exists a single evaluative standard that rationalizes all sets of index numbers that could be generated by an index number system, then the index number system can be called 'exogenous'. This term is appropriate as the index numbers generated by the index number system can be viewed as corresponding to a single evaluative perspective which is invariant to the data to which the index number system is applied, and fixed ex-ante. In contrast, if there does not exist a single evaluative standard that rationalizes all the sets of index numbers that are generated by an index number system, then the index number system can properly be called 'endogenous' as the evaluative standard that rationalizes the index numbers must vary with the data to which the index number system is applied, and so the evaluative standard cannot be viewed as having been fixed. Given the assumption that the index numbers generated by an index number system are always meaningful (and therefore rationalizable), these are the only two possibilities: the index number system is exogenous or it is endogenous.

The theorem that is proved below establishes a relation between explicit or implicit evaluative judgments and the properties of an index number system. It demonstrates that certain potentially desirable properties can be satisfied if and only if the index number system is exogenous (i.e. the index numbers that it generates can always be rationalized by a single fixed evaluative standard). Therefore, the choice of an index number system from the class that possesses the identified potentially desirable properties is equivalent to the adoption of a fixed evaluative standard. Controversies concerning the appropriate choice of an index number system from among those that satisfy the identified properties can be viewed as equivalent to controversies concerning the appropriate choice of evaluative standard.

As noted above, if a rationalizable index number system is not exogenous then it is endogenous. If an index number system is endogenous, then not all sets of index numbers generated by it can be rationalized by the same evaluative standard, although they can all be rationalized by *some* evaluative standard. The evaluative standard that rationalizes the index numbers generated by the data may be fixed at most locally (rather than globally) in the sense that as the data changes, the evaluative standard used to rationalize the index numbers generated by the data may also be required to change. The index numbers generated by applying the index number system to a set of data can be rationalized by an evaluative standard which permits meaningful internal comparison of the units (such as points in time or space) to which the elements of the data pertain. However, meaningful external comparisons across all of the points in space and time to which distinct sets of such data correspond (in each of which the index numbers are rationalized by possibly disjoint sets of evaluative standards) are not feasible without formulating a meta-evaluative criterion that permits comparative judgments concerning the relation in which distinct evaluative standards stand to one another. Such a meta-evaluative criterion must itself be justified by reference to an underlying normative perspective.

An influential class of endogenous index number systems (indeed that which is the almost exclusive focus of recent literature on index numbers) is defined in the so-called "economic approach" to index numbers which has featured the requirement that the index numbers chosen be rationalized by an evaluative standard that is completely deferential to the subjective preferences imputed to agents. Since identifying a single evaluative standard that fully defers to distinct agents simultaneously is not a coherent concept if the agents differ, it has been assumed in this approach that agents can be interpreted as sharing identical homothetic preferences. Moreover, it has been assumed that agents' observed consumption choices can be interpreted as reflecting the constrained maximization of a utility function of a specific functional form. Given these assumptions, it may be possible to show that a particular index number system is "exact" for the specified functional form of utility function in the sense that the index numbers to which the index number system would give rise (if it were applied to observed consumption data generated by utility maximizing agents who share these preferences and act in responseto the observed prices) are rationalized by an arbitrarily chosen level set of the specified utility function.[9] If the assumptions required for the counterfactual "exactness" result hold (namely, that the data is generated by utility maximizing agents who possess fixed common homothetic preferences of the presumed functional form) then index numbers generated by an exact index number system will always be rationalized by a fixed evaluative standard (namely an

[9] See e.g. Diewert (op cit), Konűs (op cit), Buyshgens (op cit).

arbitrarily chosen level set from the shared preference map). The index number system will therefore function on the restricted domain of data defined by these assumptions as if it is exogenous and it will possess the properties of exogenous index number systems on this domain.

However, if the assumptions required for the counterfactual "exactness" result do not hold, then the index numbers generated by the index number system will not necessarily be rationalizable by a fixed evaluative standard (if they are rationalizable at all). Disjoint sets of evaluative standards may be required to rationalize the index numbers produced by applying the index number system to distinct sets of data. If the index numbers produced by applying an "exact" index number system to given data are rationalizable, they can be given a meaningful counterfactual interpretation (such as, in the case of a price index, the relative cost of achieving a level set at different points in time or space). However, the index numbers produced by applying the index number system to *distinct* sets of data may no longer be rationalizable by the *same* level set, and therefore may not be possible to relate meaningfully without a meta-evaluative criterion which identifes the *relation* in which given level sets stand to one another.[10] This problem is not unique to the subjective preference centered "economic approach", but rather applies to all approaches to rationalizing index numbers in which the evaluative standard used to rationalize them is in any way sensitive to the preferences attributed to agents (as inferred on the basis of their observed choices).

Meaningful (i.e. rationalizable) index number systems must be either exogenous or endogenous. The choice of an exogenous index number system is equivalent to the choice of a fixed evaluative perspective. The choice of an endogenous index number system is also equivalent to the choice of an evaluative perspective, although more subtly: It is necessary to appeal to a normative standpoint in order to justify the substitution of one rationalizing evaluative standard for another as the data change. In particular, it is necessary to identify a criterion for meta-evaluation that permits comparisons between distinct evaluative standards in order to make comparisons over an unrestricted domain of possible data (since distinct evaluative standards may be necessary to rationalize the sets of index numbers generated at distinct data states). Since the construction and use of meaningful index numbers requires the adoption of either an endogenous or an exogenous approach, it is an activity that inescapably demands implicit or explicit normative judgments. These normative judgments necessarily depend on the purpose for which the index number is intended to be used.[11]

[10] An example of such a 'meta-evaluative criterion' is a criterion of interpersonal comparison, which identifies level sets from distinct preference maps as corresponding to a common level of achievement of welfare.

[11] The argument of this essay therefore supports that strand of reasoning about index nubmers which insists that

The exogeneity theorem proved in the paper demonstrates that index number systems that possess certain generally desirable properties (associated with meaningfulness, relevance and consistency) necessarily admit of a fixed welfare interpretation. If an index number system that does not permit a fixed welfare interpretation is adopted, then it will not be possible to undertake meaningful comparisons across all possible units of comparison without adopting a meta-evaluative criterion based on appropriate welfare judgments. In this sense, normative judgments are inescapable in the construction and use of index numbers.

In the course of laying the groundwork for the theorem, we identify a set of bounds on the index numbers that an index number system produces (the "generalized within bounds" requirement) which define necessary and sufficient conditions for these index numbers to be rationalizable (i.e. to possess a welfare interpretation). We focus on multilateral index number systems, by which we we refer to index number systems which produce index numbers that are used to compare multiple (three or more) units of comparison. It may be shown that multilateral index number systems that are widely in use violate this and other requirements of the exogeneity theorem. As a result, the index numbers they produce do not admit of a welfare interpretation which is invariant to the data to which they are applied, and indeed may not admit of any welfare interpretation at all. The exogeneity theorem therefore establishes that the rationale for "exact" index number systems is much less compelling than widely believed.

Our focus on multilateral index number systems does not constitute a significant limitation. We shall focus on multilateral index numbers because almost all interesting purposes to which index numbers can be put require multilateral index numbers. The special case of bilateral index numbers is so specialized that it is for all practical purposes irrelevant. Such units of comparison might be spatial jurisdictions (such as countries), temporal units (such as years) or units in space-time (such as countries in individual years).[12] Brief reflection suffices to create the recognition that addressing

there is no single index number system that is best for all purposes, but rather that the appropriate choice of index numbers depends on the purpose for which it is intended that they be used (For statements of this view see e.g. Afriat, S. (1977), *The Price Index*, Cambridge: Cambridge Univ. Press; Deaton, A. (1981), "Introduction" to "Part two", in *Essays in the Theory and Measurement of Consumer Behaviour*, Cambridge: Cambridge Univ. Press; Rogoff, K. (1996) "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, Vol. XXXIV; Samuelson, Paul A & Swamy, S, 1974. "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis," *American Economic Review,* vol. 64(4), pages 566-93.

[12] The first major attempt to create a comprehensive space-time tableau of economic aggregrates occurred in the 1960s. The International Comparison Program (ICP) has worked to calculate multilateral spatial index numbers since its inception in 1967. Under the stewardship of the United Nations and subsequently the World Bank, the ICP has worked with an increasing number of national statistical agencies to collect data on prices of and expenditures on an increasing number of goods. The first benchmark survey was conducted in 1970; subsequent surveys have taken place roughly every five years. The data collected by the ICP has served as the basis for the creation of the Penn World

almost any practical problem entails the use of multilateral index numbers. Bilateral index numbers often fail to be transitive, as a result of which they do not offer an unambiguous basis for comparison between entities, and are thus inappropriate for much applied work. This is a basic problem that is well known to index number theorists but not to all applied researchers. Those readers who are unfamiliar with this problem may consult the first appendix for an illustrative example. [13]

In what follows, we specialize to the case of cost of living indices in order to fix ideas, although there is no part of our argument which is specific to this choice. The extension of our results to the cases of production price indices and output quantity indices is straightforward[14].

4 Index Number Systems: Mappings from Data States to Index Numbers

The first notion we must formalize defines the information employed by the index number system. Note that we are abstracting from aggregation problems internal to individual objects of comparison (such as countries or years) and categories of commodities which must be solved in order to produce this data. For example, "bread" comes in different varieties which have different prices, and its price may vary between urban and rural areas or between seasons. For purposes of this investigation, we

Tables, certainly the most influential practical attempt to create a space-time tableau of economic aggregates. The availability of this new data, and the computer power to analyze it, has had a great impact. As economic researchers have undertaken econometric exercises drawing inferences based on intertemporal and interspatial variation, economic theorists have reason to renew their interest in the theory of index numbers applicable to multiple units of comparison .

[13] Such an example can only begin to capture the problems involved in real comparisons. The "raw" data derived from a single benchmark survey of the International Comparison Programme is a very long list of price and expenditure observations. Afterwards, a herculean effort is marshalled to transform this raw data into an estimated price structure and consumption pattern for a common set of aggregrate commodities (such as clothing, transportation, etc...) for each country in the survey. In addition to the practical issues of missing data and measurement error, there is also a great deal of "scalarization" necessary to get even to this stage, such as aggregation of observations collected in different regions and in different seasons, aggregation of observations for items actually consumed, such as "apples" into larger internationally comparable aggregates (such as "fruits and vegetables"). In other words, an initial index number problem must be solved at a detailed level in order to assemble internationally comparable price and quantity observations for commodities in order that the data needed to solve the larger problem of international price and output comparisons can be assembled. The data used in such an exercise consists of a pair of vectors (summarizing the price structure and consumption pattern) for each country. If there are only 25 aggregate commodities, and 200 countries in the world, the final output will be 200 pairs of 25-dimenisonal vectors, an ordered list of 10,000 real numbers.

[14] This is so if price and cost schedules are linear, in which case the results of the paper are entirely unchanged. A full extension of the theory to quantity indices requires the addition of a cardinalization requirement, such as that the quantity index be homogenous of degree one.

As Fisher and Shell (1998, op cit) point out it may be important to take account of the non-linearity of price and cost schedules, and their dependence on economic structure, when developing a production price index. In that case, further development of the standard theory presented here will be required.

consider these "antecedent" index number problems to have been solved in order that we can pose and consider the problem at hand[15].

Definition 1 *A context of economic comparison consists of the following elements:*

i.) A set $\mathbf{N}$ of n distinct objects of comparison, the elements of which are indexed by $i \in (1, ..n)$

ii.) A set $\mathbf{G}$ of g distinct commodities, common to all $i \in \mathbf{N}$

iii.) A "price structure", $\overline{p_i} \in \mathbb{R}^g_{++}$ for each $i \in \mathbf{N}$

iv.) A "consumption pattern", $\overline{q}_i \in \mathbb{R}^g_{+}$ for each $i \in \mathbf{N}$

$\mathbf{N}$ is a comparison set consisting of the objects of comparison (such as spatial or temporal entities) that we wish to compare. Prices may or may not be expressed in distinct currency units, and consumption is defined in terms of some standard quantity units, possibly specific to each commodity but common to the distinct objects of comparison. We assume throughout that there are at least three objects of comparison.

Definition 2 *A data state (d) is a fixed assignment of price structure and consumption pattern to each* $i \in \mathbf{N}$

$$d = \{\overline{p}_i, \overline{q}_i\}_{i \in \mathbf{N}} \tag{1}$$

It will be convenient for us to employ the following notation to refer to specific elements of a data state:

Data state price structure:

$$d_p = \{\overline{p}_i\}_{i \in \mathbf{N}} \tag{2}$$

[15] In doing so, we do not presuppose that the antecedent index number problems are themselves resolvable without recourse to normative judgments. The 'data' of the problem are taken to be given to us, but need not have been measured in a manner that is free of judgments.

Data state consumption pattern:

$$d_q = \{\overline{q}_i\}_{i \in \mathbf{N}} \tag{3}$$

Object of comparison i data ($i \in \mathbf{N}$):

$$d_i = \{\overline{p}_i, \overline{q}_i\} \tag{4}$$

Object of comparison i price structure ($i \in \mathbf{N}$):

$$d_{p_i} = \overline{p}_i \tag{5}$$

Object of comparison i consumption pattern ($i \in \mathbf{N}$):

$$d_{q_i} = \overline{q}_i \tag{6}$$

Data in all objects of comparison excluding i ($i, k \in \mathbf{N}$):

$$d_{-i} = \{d_k\}_{k \neq i} \tag{7}$$

Data in all objects of comparison excluding both i and j ($i, j, k \in \mathbf{N}$):

$$d_{-ij} = \{d_k\}_{k \neq i, k \neq i} \tag{8}$$

Definition 3 *The set of data states, S, is the collection of all possible data states that can arise in the context of economic comparison that is being considered. We assume that:*

$$\mathbf{D} = \{d | d_{p_i} \in \mathbb{R} \ and \ d_{q_i} \in \mathbb{R}^g_+ \quad for \ each \ i \in \mathbf{N}\} \tag{9}$$

Definition 4 *An index number is a positive real number. An index number set is an n-by-n matrix of index numbers. The set of index number sets is the collection of all possible index number* *sets. We shall denote the set of index number sets by* **I**.

Definition 5 *An index number system, S, is a mapping from* **D** *to* **I**.

$$\forall d \in \mathbf{D}, \qquad S(d) \in \mathbf{I} \tag{10}$$

$$S \text{ is the index number system (the mapping)} \tag{11}$$

$$S(d) \text{ is the index number set associated with data state } d\text{, an } n-by-n \text{ matrix} \tag{12}$$

$$S_{ij}(d) \text{ is the index number associated with data state } d\text{, row } i \text{ and column } j \tag{13}$$

$$i \text{ is referred to as the } \textit{focal object of comparison} \tag{14}$$

$$j \text{ is referred to as the } \textit{reference object of comparison} \tag{15}$$

The idea that we wish to capture is that an evaluator observes a "signal" (a data state) as a result of which she assigns index numbers to each pair of objects of comparison according to the method for constructing index numbers (or the index number system) that she has adopted. The signal is not known a priori. This paper is concerned with the properties of the resulting mapping, and the interpretations one may give to the index numbers that it produces.

5 Rationalizability of Index Numbers by an Evaluative Standard

In this section, we will define what it means for an index number system to admit of a meaningful welfare interpretation, in the sense that the index numbers generated by it can be rationalized by an evaluative standard. As discussed above, we will restrict our attention henceforth to price indices, in order to fix ideas, without any loss of generality. This focus will allow us to think carefully about what interpretation an index number should have, but our results can be readily extended to a quantity index, or indeed to other types of index number.

In doing so, we will draw on the concept of an *evaluative* relation instead of the more familiar concept of a *preference* relation. We will do so in order to indicate a fundamental conceptual distinction. The former refers to the rankings over commodities that an agent employs (or is assumed to employ) when choosing consumption based on observed prices . The latter refers to the rankings over commodities that an index number theorist (a possibly external evaluator) employs when constructing a price index to undertake cost of living comparisons.

An evaluator may have beliefs about the preference relation an agent possesses (which may or may not be accurate). The preference relation imputed to an agent need not correspond to the evaluative relation which the evaluator herself possesses. To take a possibly simple-minded but helpfully stark example, an evaluator might observe the consumption behavior of two drug addicts who face different relative prices between food and cocaine. The evaluator might employ one preference relation to predict or explain the behaviour of the addicts (she may in fact use a distinct subjective preference relation for each addict) but a very different evaluative relation to assess the addicts' relative well being, or to measure the relative cost of living in the two price environments. The evaluator can make *both* empirical and normative statements; each such statement will derive from implicit scalarizing functions, but the approach to scalarization used in the two exercises may be quite distinct, relating in the first case to a predictive or explanatory purpose and in the second case to a normative purpose.

It may be argued that the two types of relations - those attributed to the agent and those employed by the evaluator in assessments - should correspond, and in particular that an evaluator's judgments should conform to the subjective preferences of the agent. This is in itself a normative position, and arguably sometimes a very compelling one. Indeed, even if we put aside extreme illustrative examples such as those involving addicts or small children, it will generally be reasonable to consider what agents themselves value when assessing their actual and counterfactual well-being. An evaluative relation that is wholly insensitive to the preferences of the agents under consideration will very likely be inappropriate for applied purposes such as comparisons of the cost of living. The issue is not whether there should be some relation between the two types of judgments. It is uncontroversial to propose that there ought to be such a relation. Rather, the question is whether the two types of judgments should be constrained to be identical in all circumstances.[16] We will adopt the point of view that evaluative relations and preference relations can differ. Thus we will clearly distinguish between these two concepts throughout the paper. This approach is a straightforward generalization of the "economic approach" to index number theory and accommodates it as a special case. We hope to show in what follows that this generalization opens fruitful avenues of theoretical research and practical application. Accordingly, we propose the following definitions:

Definition 6 *An evaluative relation, $\geq_E$, is a binary relation over commodities which satisfies the following three properties $\forall\ q^0, q^1, q^2 \in \mathbb{R}^g_+$:*

$$\text{either } q^0 \geq_E q^1 \text{ or } q^1 \geq_E q^0 \text{ or both} \qquad \text{completeness} \tag{16}$$

[16] A rather forceful normative position – that the evaluative relation implicitly or explicitly employed when constructing index numbers should be identical to the preference relation possessed by the agents being compared – is adopted in what is commonly referred to as the "economic approach" to index number theory. As noted earlier, this description of the approach may be misleadingly broad, in view of the specificity of the normative perspective that serves as its foundation.

Gottfried Haberler argued this position insistently: "Die Wissenschaft macht sich einer Grenzuberschreitung schuldig, sie fallt ein Werturteil wenn sie die Wirtschaftsubjekte belehren will welches von zwei Naturaleinkommen das 'grossere' Realeinkommen enthalt. Daruber zu entscheiden, welches vorzuziehen ist, sind einzig und allein die Wirtschafter selbst berufen." ("Science is guilty of trespassing beyond its necessary limits–that is to say, it is delivering a judgment of value–if it attempts to lay down for others which of two real incomes is the 'larger.' To decide on this, to decide which real income is to be preferred, is a task which can only be done by him who is to enjoy it–that is, by the individual as 'economic subject'"). [Translated from Haberler's *Der Sinn der Indexzahlen* and cited in Lionel Robbins' *An Essay on the Nature and Significance of Economic Science* (Second Edition, Revised and Enlarged, 1935, London: MacMillan and Co., Ltd)].

If one accepts the view that index numbers ought to reflect "the preferences possessed by the agents", a great number of technical issues immediately present themselves. For example, implicit in the idea of deference to preferences is the idea either that there is a single set of preferences to which to defer or that many alternative sets of index numbers must be entertained. If the economic units being compared are comprised of many agents (possibly millions in the case of countries) there are rather serious problems of heterogeneity that may arise, and make it difficult or impossible to identify a representative agent within each such unit. If there is to be a single preference relation which the evaluator is to adopt as her evaluative relation and if the evaluator wishes simultaneously to defer to agents then every representative agent must have the same preferences. Further, the evaluator must choose a particular level set as the reference for index number construction. Within the framework of deference to agents, this requires either that all representative agents possess a common level of "real income" (i.e. the same level set is the appropriate one to consider for all of them), or that their shared preferences are homothetic, in which case the choice of level set is inconsequential for the index numbers generated. These and related issues are discussed further in the accompanying paper.

$$\text{if } q^0 \geq_E q^1 \text{ and } q^1 \geq_E q^2 \text{then } q^0 \geq_E q^2 \qquad \text{transitivity} \tag{17}$$

$$\text{if } q^0 \geq q^1 \text{ then } q^0 \geq_E q^1 \qquad \text{weak monotonicity} \tag{18}$$

Note that $\geq$ denotes weak vector dominance. If both $q^0 \geq q^1$ and $q^1 \geq q^0$ we will write $q^0 \ =_E q^1$.

Further,

Definition 7 *$\geq_E$ is weakly convex if $\forall\ q^0, q^1 \in \mathbb{R}^g_+$, $\forall\ \lambda \in (0,1)$:*

$$\textit{if } q^0 \ =_E q^1 \textit{then } \lambda q^0 + (1-\lambda)q^1 \geq_E q^i \qquad (i = 1, 2) \tag{19}$$

It will be important for us to consider a particular level set associated with an evaluative relation. Hence. we define:

Definition 8 *An evaluative standard, E^* associated with an evaluative relation $\geq_E$ is defined by:*

$$E^* = \{q \in \mathbb{R}^g_+ | q \ =_E q^* \text{ for some non-zero bundle } q^* \in \mathbb{R}^g_+\} \tag{20}$$

If the evaluative relation is weakly convex then an evaluative standard defined in relation to it shall also be referred to as weakly convex, or simply as convex.

We shall denote the set of evaluative standards by $\mathbf{E}$ and the set of convex evaluative standards by $\mathbf{E}_c$.

Definition 9 *The minimum cost function, c_{E^*}, associated with an evaluative standard E^* is the following function of prices:*

$$c(\overline{p}, E^*) = \{\min \langle \overline{p}, \overline{q} \rangle \, | \overline{q} \in E^*\} \tag{21}$$

The evaluative relation is weakly monotonic, so the minimum cost function is well-defined. The evaluative standard excludes the zero bundle, so the minimum cost function is strictly positive. We assume here and in what follows that prices are linear, for simplicity, although the argument can, *mutatis mutandi*, withstand modification of this standard assumption.[17]

Definition 10 *S is locally normatively rationalizable (LNR) at data state d if:*

$$\exists E_d^* \in \mathbf{E} \qquad s.t. \qquad S_{xy}(d) = \frac{c(d_{p_x}, E^*)}{c(d_{p_y}, E^*)} \qquad\qquad \forall x, y \in \mathbf{N} \tag{22}$$

S is *locally normatively rationalizable by a convex standard* (LNRC) at d if $E_d^* \in \mathbf{E}_c$.

Note that the rationalizing evaluative standard need not be unique. For a finite number of objects of comparison, it will generally be possible to find an infinite number of standards that rationalize the index numbers that are generated at a particular data state. LNR simply means that some welfare conception (evaluative standard) exists that can rationalize a particular index number set. It is natural to ask if such a conception exists for all possible index number sets. Hence, we can define:

[17] This is not to say that the exercise is trivial. That it is not is shown by Fisher and Shell (1998), op cit.

Definition 11 *S is globally normatively rationalizable (GNR) if* $\forall\ d \in \mathbf{D}$

$$\exists E_d^* \in \mathbf{E}\ s.t. \qquad S_{xy}(d) = \frac{c(d_{p_x}, E^*)}{c(d_{p_y}, E^*)} \qquad \forall x, y \in \mathbf{N} \tag{23}$$

S is GNR for a convex evaluative standard if $E_d^* \in \mathbf{E}_c\ \forall\ d \in \mathbf{D}$.

Note that the rationalizing evaluative standard can vary from one data state to the next. Given that the rationalizing standard at each data state may be non-unique, it is most accurate for us to visualize *sets* of rationalizing standards associated with each data state. GNR requires that each associated set of rationalizing standard be non-empty. Finally, it is natural for us to ask if their *intersection* is non-empty, in which case a single evaluative standard can be found which rationalizes all sets of index numbers that may be generated by the index number system:

Definition 12 *S is exogenous if the following holds:*

$$\exists E^* \in \mathbf{E}\ s.t.\ S_{xy}(d) = \frac{c(d_{p_x}, E^*)}{c(d_{p_y}, E^*)} \qquad \forall x, y \in \mathbf{N}, \forall d \in \mathbf{D} \tag{24}$$

S is exogenous for a convex evaluative standard if $\forall\ d \in \mathbf{D},\ \ E^* \in \mathbf{E}_c$.

If S is GNR but not exogenous, S is *endogenous*.

LNR requires that a given set of index numbers (generated at a particular data state) possesses a welfare interpretation. GNR requires that all sets of index numbers that may be generated possess a welfare interpretation. Exogeneity requires that all sets of index numbers that may be generated possess a common welfare interpretation.

We have now made rigorous the notion of what it means for an index number system to admit of a welfare interpretation. However, as we shall see, virtually all index number systems commonly in use fail to satisfy GNR, let alone exogeneity. If they admit of a welfare interpretation, they do so only on a restricted domain. However, if an index number theorist wanted an index number system which was exogenous, she could easily find one. As discussed below, by starting with a fixed evaluative standard, she could simply construct the corresponding index number system.

5.1 Examples of Exogenous Index Number Systems

Let us consider some examples of exogenous index number systems. This will help motivate the discussion of the remaining properties with which we are concerned.

Example 1 *The Fixed-Bundle Index Number System*

Fix a non-zero bundle $q^* \in \mathbb{R}^g_+$ and define an associated index number system, S^{q^*}, which conveys the relative cost in each object of comparison of purchasing this fixed bundle. In particular:

Definition 13 *The index number system, S^{q^*}, associated with q^*:*

$$S^{q^*}_{xy}(d) = \frac{\langle q^*, d_{p_x} \rangle}{\langle q^*, d_{py} \rangle} \tag{25}$$

This index number system has a clear interpretation; it conveys the relative cost of purchasing bundle q^* at the prices faced respectively in objects of comparison x and y. This may be what many lay people have in mind when they think about a price index. It is certainly a natural and readily understandable approach. The fixed basket of goods may be chosen in order to be in some way "representative". No matter how one specifies the basket, the major drawback is that it is assumed that each agent ought to be able to purchase an identical basket even though the price environments that they face differ. A more flexible approach would allow for the possibility of substitution between goods in response to variant prices. Such substitution is captured by a simple generalization of the bundle approach:

Example 2 *The Fixed-Standard Index Number System*

Fix an evaluative standard $E^* \in \mathbf{E}$ and define an associated index number system, S^{E^*}, which conveys the relative minimum cost in each object of comparison of achieving this fixed standard. In particular:

Definition 14 *The index number system, S^{E^*}, associated with E^*:*

$$S_{xy}^{E^*}(d) = \frac{c(d_{p_x}, E^*)}{c(d_{p_y}, E^*)} \tag{26}$$

It is obvious that S^{E^*} is exogenous. It is also a straightforward generalization of the fixed-bundle index number system. Instead of requiring each object of comparison to be assigned an income sufficient to purchase an identical bundle, we now require each object of comparison to be assigned an income sufficient to achieve an identical "standard of living" (understood as attaining the evaluative standard E^*) by consuming the commodity bundle that permits this at least cost in each object of comparison. We have simply extended the notion of sameness from that of achieving an identical bundle to achieving an identical "standard of living" as defined by an evaluative standard. In fact, the fixed-bundle index number system is a special case of the fixed-standard index number system, in which the evaluative relation corresponds to Leontief preferences and q^* is a commodity bundle that is a least-cost means of achieving an evaluative standard associated with this evaluative relation.

Note that any exogenous index number system trivially satisfies GNR:

Lemma 1 *If S is exogenous, S is GNR.*

Proof. The exogenous standard E^* normatively rationalizes every data state. ■ ■

There are a number of other properties that an exogenous index number system satisfies. Let us consider three of particular importance.

5.2 Transitivity

The first property we will consider is transitivity, which requires the equivalence of direct and indirect comparisons.

Definition 15 *S is transitive if the following equality holds:*

$$S_{xy}(d) = \frac{Sxz(d)}{Syz(d)} \qquad \forall x, y, z \in \mathbf{N}, \forall d \in \mathbf{D} \tag{27}$$

This is a standard property demanded of almost every index number system employed in practice, as in its absence comparisons between objects become potentially ambiguous. The term transitivity is used in a number of related but separate ways in the literature; we provide a brief treatment of the relation between transitivity and related concepts in the second appendix so as to resolve any ambiguities that may arise.

Lemma 2 *If S is exogenous, then S is transitive.*

Proof. This follows immediately from the "quotient separable" form of an exogenous index number system:

$$S_{xy}^{E^*}(d) = \frac{c(d_{p_x}, E^*)}{c(d_{p_y}, E^*)} = \frac{\frac{c(d_{p_x}, E^*)}{c(d_{p_z}, E^*)}}{\frac{c(d_{p_x}, E^*)}{c(d_{p_z}, E^*)}} = \frac{S^{E^*}xz(d)}{S^{E^*}yz(d)} \tag{28}$$

■

This manipulation is valid since the cost function is strictly positive. ■

5.3 Anonymity

An important idea is that the only information about objects of comparison that is relevant to determining a set of index numbers is contained in the data state and not in the identities of the objects of comparison possessing specific data. This is a standard property demanded of index number systems, along with transitivity. In particular:

Definition 16 *A pair-wise permutation operator, η_{xy} is a mapping from **D** to **D**:*

$$\eta_{xy}(d) = \{d' | d'_x = d_y, d'_y = d_x, and d'_{-xy} = d_{-xy}\} \tag{29}$$

Definition 17 *S is anonymous if the following equality holds:*

$$S_{xz}(d) = S_{yz}(d') \text{ where } d' = \eta_{xy}(d) \qquad \forall x, y, z \in \mathbf{N}, \forall d \in \mathbf{D} \tag{30}$$

We may note that:

Lemma 3 *If S is exogenous, then S is anonymous.*

Proof. This follows immediately from the "quotient separable" form of an exogenous index number system: ■

$$S_{xz}(d) = \frac{c(d_{p_x}, E^*)}{c(d_{pz}, E^*)} = \frac{c(d'_{p_y}, E^*)}{c(d'_{p_z}, E^*)} = S_{yz}(d') \quad \blacksquare \tag{31}$$

Anonymity is a seemingly minimal property which in actuality imposes significant requirements. Anonymity entails in the present context that the only information we are required to know about distinct objects of comparison to determine their relative cost of living concerns their respective price structures and consumption patterns. This is a severe restriction on the information that may be used by the evaluator, who is prohibited from considering any other information about objects of comparison. Note that the term object of comparison has been used here as an abstract term which may denote individual persons or groups of persons. Anonymity would require that the cost of living assessment for a particular person should be independent of features of the person other than the prices they face and the quantities they consume, and corresponding information for other persons. The information set to which cost of living assessments are permitted to be responsive is accordingly restricted, and *inter alia* a person's age, gender, health status, physical constitution, personal disposition, or occupation are to be deemed irrelevant to the assessment. Implicitly, this requirement rules out the possibility of adjusting the index numbers associated with different persons (or groups of persons) so as to reflect the fact that they transform commodities into well-being in different ways. This requirement is therefore quite stringent and its wisdom may reasonably be questioned.

Transitivity and anonymity are two standard properties of index number systems. Both requirements appear conceptually minimal, but they impose a great deal of structure on the index number problem, and significantly restrict the class of index number systems that an evaluator is permitted to employ.

5.4 Bilateral Comparison Focus

Now that we have considered two properties that are satisfied by virtually all index number systems generally employed, let us now turn our attention to a property that is almost categorically violated.

Definition 18 *S satisfies bilateral comparison focus (BCF) if the following holds:* $\forall x, y \in \mathbf{N}, \forall d \in \mathbf{D}$

$$S_{xy}(d_x, d_y, d_{-xy}) = S_{xy}(d_x, d_y, d'_{-xy}) = S_{xy}(d_x, d_y) \tag{32}$$

Lemma 4 *If S is exogenous, then S satisfies BCF.*

Proof. By inspection, S depends only on data (and in particular, prices) in objects of comparison x and y. ■ ■

Bilateral comparison focus requires that an assessment of the relative cost of living between two objects of comparison should depend only on the data of those two objects of comparison; it should not depend on information from third entities. Why should the relative cost of living between the US and Zambia vary with Japan's consumption? There are at least two perfectly good reasons to violate bilateral comparison focus which are immediately apparent. Firstly, the evaluator might adopt a 'relativist' rather than an absolutist point of view, in which the well-being of an individual is influenced by the consumption of others, in which case the assessments of the cost of living in any one object of comparison ought very properly to depend on information from all the other objects of comparison. Secondly, the evaluator might use information from all objects of comparison to form judgments as to what evaluative standard to employ for living standards comparisons. For example, if the evaluator believes that agents in all objects of comparison share common homothetic preferences, then she may use observed consumption in all the objects of comparison as a signal with which to form judgments as to what these preferences are, in order to isolate a level-set in the commodity space for use in cost of living assessments[18]. However, if the evaluator is interested in an absolute standard of achievement (for example, the resources required to achieve a nutritional norm) then bilateral comparison focus appears to be a very reasonable requirement. Bilateral comparison focus is a strong version of the "characteristicity" requirement that has been widely discussed in the index number literature[19]. Whereas the characteristicity requirement is that the index number between two objects of comparison should reflect the data only in those two objects of comparison *to the greatest extent possible*, bilateral comparison focus is the requirement that the index number must be influenced *only* by the data in the objects of comparison concerned. To the best of our knowledge, characteristicity has never been formally defined.

Clearly, BCF is not desirable in every context. However, BCF is a natural point of departure in the assessment of index numbers; deviations from BCF should be carefully considered and adequately justified. The question is not merely if a bilateral comparison should vary with third object data, but rather *how exactly* the comparison should vary with third object data, if it should vary at all. If an index number system violates BCF in an inappropriate manner, or if the variation of index numbers in response to third object data is so complex that one cannot determine whether the index numbers that are assigned vary *appropriately* with third object data, then this violation of BCF may be a serious issue. From a practical perspective, is would mean that the system is not robust to sources of *irrelevant* variation. Indeed, although the term bilateral comparison focus will be used throughout this paper, the concept to which it refers is intimately related to the widely employed idea of "independence of irrelevant alternatives". We choose not to use the latter expression because the "alternatives" may not always be irrelevant, as just discussed. However, the "alternatives" may indeed be irrelevant, in light of the task of comparison which is being undertaken, in which case BCF is a property with which we should be greatly concerned.

5.5 The Exogeneity Theorem (Part I)

[18] On the signal extraction interpretation of existing index number systems, see Reddy, S. and B. Plener, "The Choice of Index Number: Part II", forthcoming.

[19] See e.g. Drechsler, L. (1973), "Weighting of Index Numbers in Multilateral International Comparisons", *Review of Income and Wealth*, Vol. 19, pp. 17-34.

Theorem 1 *If S is exogenous for a convex evaluative standard, then S is GNR by a monotonic and convex evaluative standard, BCF, transitive and anonymous.*

Proof. This follows from Lemmas 1, 2, 3 and 4. ■

A more difficult question is whether an index number system must be exogenous in order to satisfy these four properties. We will address this question directly, in the affirmative, soon. First, we must develop a framework that will enable us to geometrically represent sets of index numbers.

6 A Diagrammatic Representation of Index Numbers

We have seen how an explicit evaluative standard gives rise to index number systems with potentially desirable properties. Is it possible to achieve those properties without an evaluative standard? Or is it the case that behind every "good" index number system there is a rationalizing evaluative standard? If so, can we determine what this standard looks like? It turns out that the answers to these questions are intimately related to the "geometry" of index number systems. We will now introduce the concepts necessary to make this notion concrete.

Consider the following thought experiment. Given a data state, d, and an index number system, S, fix a reference object of comparison, y. Give y one unit of its local currency. Give object of comparison x $S_{xy}(d)$ units of its local currency. Now object of comparison x has an endowed level of income, $S_{xy}(d)$, and a fixed price structure, d_{p_x}. What goods can x afford, with that income, and those prices? We will call that set of affordable bundles the *feasibility set*, $F_{xy}(d)$. We will use the term "feasibility set" instead of "budget set" because the income here has been arbitrarily assigned, and will generally not correspond to the actual income of object of comparison x. In this exercise, object of comparison y merely serves an accounting role, determining the income assigned to object of comparison x. While $S_{xy}(d)$ is denominated in object of comparison x local currency units per object of comparison y local currency units, we will multiply by one unit of object of comparison y currency to get an income expressed in object of comparison x currency. This is simply a normalization procedure. Other normalization procedures are also available to us, but this one will serve our purpose fully.

Definition 19 *The feasibility set $F_{xy}(d)$ is the collection of bundles that can be afforded in object of comparison x at prices d_{p_x} with income $S_{xy}(d)$:*

$$F_{xy}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, d_{p_x} \rangle \leq S_{xy}(d)\} \tag{33}$$

Definition 20 *The exterior of a feasibility set, $extF_{xy}(d)$, is the collection of bundles that cannot be afforded in object of comparison x at prices d_{p_x} without exhausting the assigned income $S_{xy}(d)$:*

$$extF_{xy}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, d_{p_x} \rangle \geq S_{xy}(d)\} \tag{34}$$

Definition 21 *The interior of a feasibility set, $intF_{xy}(d)$, is the collection of bundles that can be afforded in object of comparison x at prices d_{p_x} without exhausting the assigned income $S_{xy}(d)$:*

$$intF_{xy}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, d_{p_x} \rangle < S_{xy}(d)\} \tag{35}$$

For a fixed choice of index number system, a given feasibility set is identified by the triplet of parameters, $x \in \mathbf{N}$, $y \in \mathbf{N}$ and $d \in \mathbf{D}$.

Let $\beta = (x, y, d)$, let $\mathbf{B} = \{\beta \mid x \in \mathbf{N},\ y \in \mathbf{N} \text{ and } d \in \mathbf{D}\}$ and let B denote an arbitrary subset of $\mathbf{B}$. We will call B an identifier set.

Definition 22 *The feasibility profile $\mathbf{F}_B$ associated with index number system S and identifier set B is the following collection of feasibility sets:*

$$\mathbf{F}_B = \{F_\beta = F_{xy}(d) | \beta = (x, y, d) \in B\} \tag{36}$$

It will be convenient for us to employ the following notation in reference to feasibility profiles:

Feasibility profile $\mathbf{F}_B$ with feasibility set $F_{\overline{\beta}}$ excluded:

$$\mathbf{F}_{B-\overline{\beta}} = \{F_\beta | \beta \in B, \beta \neq \overline{\beta}\} \tag{37}$$

The exterior of feasibility profile $\mathbf{F}_B$:

$$ext\mathbf{F}_B = \bigcap_B extF_\beta \tag{38}$$

The interior of feasibility profile $\mathbf{F}_B$:

$$int\mathbf{F}_B = \bigcup_B intF_\beta \tag{39}$$

These concepts will give us some powerful tools with which to think about index numbers. The most important tool of all will be the concept of generalized within bounds, to be defined in the next section.

6.1 The Generalized Within Bounds Requirement and Rationalizability by an Evaluative Standard

The Generalized Within Bounds requirement specifies a geometric property of feasibility profiles which is intimately connected to (indeed, necessary and sufficient for) the existence of a welfare interpretation of index numbers. We will first provide an abstract definition and then proceed to interpret it more concretely:

Definition 23 $\mathbf{F}_B$ *satisfies generalized within bounds (GWB) if:*

$$\forall \beta \in B, \qquad F_\beta \cap ext\mathbf{F}_{B-\beta} \neq 0 \tag{40}$$

If there are two feasibility sets, GWB requires that one not lie strictly within the interior of the other. If this occurred, then anyone with monotonic preferences would be strictly better off with the more encompassing feasibility set. When GWB is satisfied by a feasibility profile consisting of two feasibility sets, we say that it satisfies the "within bounds" requirement or more specifically that the assigned index number S_{xy} is "within bounds". If there are multiple feasibility sets, GWB requires that one not lie within the union of the others' interiors. GWB rules out this monotonic domination, ensuring that each feasibility set appears the most attractive given *some* monotonic evaluative relation. GWB states that everybody should be able to consume some bundle of which nobody else can consume strictly more.

We are interested in this requirement, because we want to rationalize the feasibility profile with an evaluative standard, a set of bundles deemed equivalent by some monotonic evaluative relation. If GWB were violated and monotonicity of the evaluative relation is demanded, then it could never be plausible to suggest that an agent commanding the strictly contained feasibility set is as well off as other agents. An agent who possesses a feasibility set other than the contained one must be deemed to be strictly better off (assuming that the same evaluative relation is to be used to assess living standards in all instances). This would be inconsistent with the premise that the index numbers reflect the relative cost of achieving some common level of achievement of the same evaluative relation. Figures 1 and 2 respectively graphically represent feasibility profiles for which GWB is satisfied and for which it is not satisfied. An important observation about GWB is that it is a property of the index numbers assigned to countries together with their price vectors, and in no way depends on actual consumption patterns in the countries concerned, in contrast with other requirements on index numbers which are prominent in the literature (in particular, the so-called Afriat bounds).

To make this idea rigorous, we must first define what we mean by an evaluative standard that rationalizes a feasibility profile.

Definition 24 $E^* \in \mathbf{E}$ *rationalizes* $\mathbf{F}_B$ *if* $\forall\ \beta = (x, y, d) \in B$

$$S_{xy}(d) = c(d_{p_x}, E^*) \tag{41}$$

We shall denote this relationship between $E^*$and $\mathbf{F}_B$ by $E^*R\ \mathbf{F}_B$. Note that the evaluative standard is rationalizing not the behaviour of agents, but the index numbers assigned by the evaluator. The standard gives the index numbers a welfare interpretation; the relative minimum cost of achieving E^* at different prices. One way to understand the conceptual content of GWB is to consider its implications. An immediate consequence of GWB is that two feasibility sets with the same price structures which belong to a feasibility profile that satisfies GWB must have the same assigned income:

Lemma 5 *If* $F_1, F_2 \in \mathbf{F}$ *where:*

$$F_1 = \{\overline{q} \in \mathbb{R}^g_+ |\, \langle \overline{q}, p\rangle \leq Y^1)\} \tag{42}$$

$$F_2 = \{\overline{q} \in \mathbb{R}^g_+ |\, \langle \overline{q}, p\rangle \leq Y^2)\} \text{ where } Y^1 \neq Y^2 \tag{43}$$

Then $\mathbf{F}$ violates GWB.

Proof. Suppose, without loss of generality, that $Y^1 > Y^2$. It then follows that:

$$q \in F_1 \Longrightarrow \langle \overline{q}, p\rangle \leq Y^1 \Longrightarrow \langle \overline{q}, p\rangle < Y^2 \longrightarrow q \notin extF_2 \tag{44}$$

■

This implies that $\mathbf{F}$ violates GWB, i.e.:

$$F_1 \cap extF_2 = 0 \ \blacksquare \tag{45}$$

The next lemma articulates the connection between GWB and rationalizability. A feasibility profile can only be rationalized by an evaluative standard if the profile satisfies GWB.

Lemma 6 The Rationalizability Lemma

If $\mathbf{F}_B$ *violates GWB, then* $\nexists\, E^* \in E$ *s.t.* $E^* R\, \mathbf{F}_B$.

Proof. Suppose $\mathbf{F}_B$ violates GWB. Then $\exists\, \beta \in B$ s.t.:

$$F_\beta \cap ext\mathbf{F}_{B-\beta} = 0 \tag{46}$$

■

DeMorgan's Law allows us to reformulate this expression in terms of interior sets:

$$F_\beta \in int\mathbf{F}_{B-\beta} \tag{47}$$

But this means that $\forall q \in F_\beta \;\; \exists\, \beta^{'} \in B$ s.t.:

$$q \in intF_{\beta'} \tag{48}$$

But by the definition of the interior of a feasibility set, that means that $\exists\, \lambda > 1$ s.t. :

$$\lambda q \in F_{\beta'} \tag{49}$$

Putting these facts together, we get the following:

$$\forall q \in F_{\beta}, \qquad \lambda q \in F_{\beta'} \qquad \text{for some } \lambda > 1\text{and } \beta^{'} \in B \tag{50}$$

However, for any monotonic evaluative relation, we have:

$$\lambda q >_E q \qquad for\ \lambda > 1 \tag{51}$$

Therefore, there can be no standard $E^* \in E$ *s.t.* $E^* R\, \mathbf{F}_B$. If there were such a standard, that would entail that the minimum cost of attaining E^* in $\beta = (x, y, d)$ is $S_{xy}(d)$ which (by (31)) can be achieved through consuming some bundle $q^* \in F_\beta$. However, by (39), $\exists\ \beta' = (x', y', z')$ s.t. $q^* \in int$ $F_{\beta'}$which implies that $S_{x'y'}(d')$ is not the minimum cost of attaining E^* in β', contrary to (31). ■

This demonstrates that GWB is necessary for rationalizability. This result will serve a critical role in our proof of the second part of the exogeneity theorem.

7 The Exogeneity Theorem (Part II)

We will now prove the exogeneity theorem:

Theorem 2 THE EXOGENEITY THEOREM

If S is GNR, BCF, anonymous and transitive then S is exogenous for a convex evaluative standard.

7.1 Proof Strategy

Our proof strategy is straightforward. In particular, from the index numbers to which the index number system gives rise at different states, we will uncover an implicit relative cost function. From the concave relative cost function, we will trace out a 'dual' convex surface which rationalizes all the sets of index numbers that can feasibly arise. The surface which we will have constructed *is* the exogenous standard.

In order to infer an implicit relative cost function from the index number system, we will initially fix reference object of comparison data and exploit anonymity, BCF and GNR. The relationship between GWB and rationalizability will be used to demonstrate the existence and concavity of the cost function. We will then use a well-known duality theorem to infer a convex surface (of an evaluative standard) from the concave relative cost function. Finally, we will use transitivity to show that this evaluative standard rationalizes all sets of index numbers to which the index number system gives rise, thus allowing the identity of the reference object of comparison and the data assigned to it to vary freely. By doing so, we will have demonstrated that the index number system is exogenous for a convex evaluative standard.

7.2 Identifying the Implicit Relative Cost Function

In this section we shall prove some lemmas which are necessary for the proof of the theorem that we offer. We will prove them by exploiting both the "within data state" and the "between data state" invariances of index numbers that are imposed by specific properties of an index number system.

To begin, let us fix an arbitrary reference object of comparison $b \in \mathbf{N}$ and fix the data d_b^*, for some arbitrary base object of comparison:

$$d_b^* = \{d_{p_b}^*, d_{q_b}^*\} \tag{52}$$

Define an associated set of data states $\mathbf{D}_b^*$, subsuming this data:

$$\mathbf{D}_b^* = \{d \in \mathbf{D} | d_b = d_b^*\} \tag{53}$$

For index number system S, consider the associated set of index numbers $\mathbf{I}_b^*$:

$$\mathbf{I}_b^* = \{S_{xb}(d) | x \in \mathbf{N}, d \in \mathbf{D}_b^*\} \tag{54}$$

If S is BCF and anonymous, then $S_{xb}(d) \in \mathbf{I}_b^*$ depends only on the focal object of comparison's data:

Lemma 7 *If S is BCF and anonymous, $d, d' \in \mathbf{D}_b^*$ and $d_x = d'_y$ then $S_{xb}(d) = S_{yb}(d')$*

Proof. *Consider* $d'' \in \mathbf{D}_b^*$ such that $d''_x = d_x = d'_y = d''_y$. By BCF: ■

$$S_{xb}(d) = S_{xb}(d'') \tag{55}$$

Note that $d'' = \theta_{xy}(d'')$. Thus, by anonymity:

$$S_{xb}(d'') = S_{yb}(d'') \tag{56}$$

Again, by BCF:

$$S_{yb}(d'') = S_{yb}(d') \tag{57}$$

Combining 55, 56 and 57:

$$S_{xb}(d) = S_{xb}(d'') = S_{yb}(d'') = S_{yb}(d') \tag{58}$$

This gives us our desired result:

$$S_{xb}(d) = S_{yb}(d') \quad \blacksquare \tag{59}$$

Thus, if S is BCF and anonymous, then $S_{xb}(d) \in \mathbf{I}_b^*$ depend only on d_x. Having fixed the reference object of comparison's data, all that matters to determining a bilateral index number is the data in the focal object of comparison, regardless of its identity. We shall now show that if S is also GNR, these index numbers depend only on prices:

Lemma 8 *If S is BCF, anonymous and GNR, $d, d' \in \mathbf{D}_b^*$ and $d_{p_x} = d'_{p_x}$ then $S_{xb}(d) = S_{xb}(d')$*

Proof. Consider $d'' \in \mathbf{D}_b^*$ such that $d''_x = d_x$ and $d''_y = d'_x$. By 7:

$$S_{xb}(d') = S_{yb}(d'') \tag{60}$$

■

Furthermore, by BCF:

$$S_{xb}(d'') = S_{xb}(d) \tag{61}$$

Now, suppose that $S_{xb}(d) \neq S_{xb}(d')$. In this case, we have the following:

$$S_{xb}(d'') = S_{xb}(d) \neq S_{xb}(d') = S_{yb}(d'') \tag{62}$$

It follows immediately that:

$$S_{xb}(d'') \neq S_{yb}(d'') \tag{63}$$

But note that $d''_{p_x} = d''_{p_y}$ by the definition of d''. By 5:

$$\{F_{xb}(d'')\}_{x\in\mathbf{N}} \text{ violates GWB.} \tag{64}$$

But then by 6, S is not rationalizable at d''. This is a contradiction, since S is GNR by assumption. Thus, our desired inequality must hold:

$$S_{xb}(d) = S_{xb}(d') \quad \blacksquare \tag{65}$$

Thus, $S_{xb}(d) \in \mathbf{I}_b^*$ depends only on the focal object of comparison's prices. This means that we can associate a function $S_{d_b^*}(p)$ with $\mathbf{I}_b^*$:

Corollary 1 *If S is BCF, anonymous and GNR then $\exists\ S_{d_b^*}(p) : \mathbb{R}^g_{++} \longrightarrow \mathbb{R}_+$ such that:*

$$S_{xb}(d) = S_{d_b^*}(d_{p_x}) \in S_{d_b^*}(p) \qquad \forall S_{xb}(d) \in \mathbf{I}_b^* \tag{66}$$

7 tells us that the index number is a function only of the focal object of comparison's data. 8 tells us that the index number is a function only of prices. It follows immediately that the index number is a function only of the focal object of comparison's prices. We now have an implicit relative cost function $S_{d_b^*}(p)$. In order to find the surface which is dual to $S_{d_b^*}(p)$ we must first ensure that $S_{d_b^*}(p)$ is weakly concave. This will follow from GWB.

Lemma 9 *If S is BCF, anonymous and GNR then $S_{d_b^*}(p)$ is weakly concave.*

Proof. $\forall \overline{p}^1, \overline{p}^2$ and $\forall\ \lambda \in (0,1)\ \ \exists\ d \in \mathbf{D}_b^*$ such that:

$$d_{p_x} = \overline{p}^1 \tag{67}$$

$\blacksquare$

$$d_{p_y} = \overline{p}^2 \tag{68}$$

$$d_{p_z} = \overline{p}^\lambda = \lambda\overline{p}^1 + (1-\lambda)\overline{p}^2 \tag{69}$$

This is always possible since we can either identify b with x, y, or z or consider it to be distinct from x, y, or z.

Consider the associated feasibility sets:

$$F_{xb}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, d_{p_x} \rangle \leq S_{xb}(d)\} \tag{70}$$

$$F_{yb}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, d_{p_y} \rangle \leq S_{yb}(d)\} \tag{71}$$

$$F_{zb}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, d_{p_z} \rangle \leq S_{zb}(d)\} \tag{72}$$

By virtue of 67, 68, 69 and 1, we can redescribe these feasibility sets:

$$F_{xb}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, \overline{p}^1 \rangle \leq S_{d_b^*}(\overline{p}^1)\} \tag{73}$$

$$F_{yb}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, \overline{p}^2 \rangle \leq S_{d_b^*}(\overline{p}^2)\} \tag{74}$$

$$F_{zb}(d) = \{\overline{q} \in \mathbb{R}^g_+ | \langle \overline{q}, \overline{p}^\lambda \rangle \leq S_{d_b^*}(\overline{p}^\lambda)\} \tag{75}$$

Now consider the feasibility profile, $\mathbf{F}_d$:

$$\mathbf{F}_d = \{F_{xb}(d), F_{yb}(d), F_{zb}(d)\} \tag{76}$$

By, GNR, d is normatively rationalizable, which means that $\mathbf{F}_d$ must satisfy GWB [by 6]. By virtue of 73, 74, 75 and the definition of GWB this means that $\exists\ \overline{q}^{\lambda} \in \mathbb{R}^g_+$ such that:

$$\left\langle \overline{q^{\lambda}}, \overline{p}^1 \right\rangle \geq S_{d_b^*}(\overline{p}^1) \tag{77}$$

$$\left\langle \overline{q^{\lambda}}, \overline{p}^2 \right\rangle \geq S_{d_b^*}(\overline{p}^2) \tag{78}$$

$$\left\langle \overline{q^{\lambda}}, \overline{p}^{\lambda} \right\rangle \leq S_{d_b^*}(\overline{p}^{\lambda}) \tag{79}$$

These inequalities guarantee the weak concavity of $S_{d_b^*}(\overline{p})$ since by combining 77 and 78.

$$\lambda S_{d_b^*}(\overline{p}^1) + (1-\lambda) S_{d_b^*}(\overline{p}^2) \leq \lambda \left\langle \overline{q^{\lambda}}, \overline{p}^1 \right\rangle + (1-\lambda) \left\langle \overline{q^{\lambda}}, \overline{p}^2 \right\rangle, \tag{80}$$

$$\lambda \left\langle \overline{q^{\lambda}}, \overline{p}^1 \right\rangle + (1-\lambda) \left\langle \overline{q^{\lambda}}, \overline{p}^2 \right\rangle = \left\langle \overline{q^{\lambda}}, \lambda\overline{p}^1 + (1-\lambda)\overline{p}^2 \right\rangle, \text{ by the linearity of the inner product.} \tag{81}$$

$$\left\langle \overline{q^{\lambda}}, \lambda\overline{p}^1 + (1-\lambda)\overline{p}^2 \right\rangle = \left\langle \overline{q^{\lambda}}, \overline{p}^{\lambda} \right\rangle \leq S_{d_b^*}(\overline{p}^{\lambda}), \text{ by the definition of } \lambda \text{ and by 79.} \tag{82}$$

Combining these three expressions, we see immediately that:

$$\lambda S_{d_b^*}(\overline{p}^1) + (1-\lambda) S_{d_b^*}(\overline{p}^2) \leq S_{d_b^*}(\overline{p}^\lambda) \tag{83}$$

Thus, $S_{d_b^*}(\overline{p})$ is weakly concave as desired. ■

7.3 Identifying an Implicit Evaluative Standard

Now that we have a weakly concave cost function associated with $\mathbf{I}_b^*$, we can use a standard duality theorem to associate an evaluative standard with $\mathbf{I}_b^*$.

Lemma 10 *If $S_{d_b^*}(\overline{p})$ is weakly concave, then $\exists\ E^* \in \mathbf{E}_c$ such that:*

$$S_{d_b^*}(\overline{p}) = c(\overline{p}, E^*) \qquad \forall \overline{p} \in \mathbb{R}_{++}^g \tag{84}$$

Proof. ■

By the Legendre transform application of the Fenchel Biconjugation [20], $S_{d_b^*}(\overline{p})$ is dual to a unique $E^* \in E^c$ s.t. $E^* R\ \mathbf{F}_B$. ■

This result implies that a single weakly convex evaluative standard rationalizes all $S_{xb}(d) \in \mathbf{I}_b^*$:

Lemma 11 *If S is BCF, anonymous and GNR, then $\exists\ E^* \in \mathbf{E}_c$ such that:*

$$S_{xb}(d) = c(d_{p_x}, E^*) \qquad \forall S_{xb}(d) \in \mathbf{I}_b^* \tag{85}$$

[20] See e.g. section 4.2, Borwein, J. and A.M. Lewis (2000), *Convex Analysis and Nonlinear Optimization: Theory and Examples*, New York: Springer. For a well-known instance of the implicit use of this duality result in economics, see "Parable and Realism in Capital Theory: The Surrogate Production Function", Paul A. Samuelson, *The Review of Economic Studies*, Vol. 29, No. 3. (June 1962), pp. 193-206.

Proof. By 1, $\exists\ S_{d_b^*}(p) : \mathbb{R}^g_{++} \longrightarrow \mathbb{R}_+$ such that: ■

$$S_{xb}(d) = S_{d_b^*}(d_{p_x}) \qquad \forall S_{xb}(d) \in \mathbf{I}_b^* \tag{86}$$

By 9, $S_{d_b^*}(d_{p_x})$ is weakly concave. Thus, by 10: $\exists\ E^* \in \mathbf{E}_c$ such that:

$$S_{d_b^*}(\overline{p}) = c(\overline{p}, E^*) \qquad \forall \overline{p} \in \mathbb{R}^g_{++} \tag{87}$$

Combining 86 and 87: $\exists\ E^* \in \mathbf{E}_c$ such that:

$$S_{xb}(d) = S_{d_b^*}(d_{p_x}) = c(d_{p_x}, E^*) \qquad \forall S_{xb}(d) \in \mathbf{I}_b^* \ \blacksquare \tag{88}$$

We have finally found an evaluative standard which rationalizes $S_{xb}(d) \in \mathbf{I}_b^*$.

Figure 3 illustrates the relation between a feasibility profile (which satisfies GWB and contains a feasibility set for each possible set of prices) and an iso-cost surface (and, implicitly, an evaluative standard).

7.4 The Exogeneity Theorem (Part II)

In order to complete the proof, we must show that the rationalizing evaluative standard we have identified is independent of the choice of base object of comparison and its data and that it therefore rationalizes all the index numbers produced by the index number system. This is true as a consequence of transitivity:

Theorem 3 The Exogeneity Theorem

If S is GNR, BCF, anonymous and transitive then S is exogenous for a convex evaluative standard.

Proof. $\blacksquare$

By 11, $\exists\, E^* \in \mathbf{E}_c$ such that:

$$S_{xb}(d) = c(d_{p_x}, E^*) \qquad \forall S_{xb}(d) \in \mathbf{I}_b^* \tag{89}$$

Define a transcription operator, Φ_z^*, in relation to some object of comparison, z, as follows:

$$\Phi_z^*(d) = \{d' | d'_z = d_b^* \text{ and } d'_{-z} = d_{-z}\} \tag{90}$$

Then the following must hold by virtue of BCF and transitivity, for arbitrary d:

$$S_{xy}(d) = S_{xy}(d') = \frac{S_{xz}(d')}{S_{yz}(d')} \qquad z \neq x, z \neq y, d' = \Phi_z^*(d) \tag{91}$$

But $d' \in \mathbf{D}_z^*$ by construction, so equation 89 holds:

$$S_{xy}(d) = S_{xy}(d') = \frac{S_{xz}(d')}{S_{yz}(d')} = \frac{c(d_{p_x}, E^*)}{c(d_{p_y}, E^*)} \qquad E^* \in \mathbf{E}_c \tag{92}$$

Since the choice of x, y and d were all arbitrary, this equality must hold universally:

$$S_{xy}(d) = \frac{c(d_{p_x}, E^*)}{c(d_{p_y}, E^*)} \qquad E^* \in \mathbf{E}_c \qquad \forall x, y \in \mathbf{N}, \forall d \in \mathbf{D} \tag{93}$$

We therefore conclude that S is exogenous for a convex evaluative standard. $\blacksquare$

7.5 An Important Extension: Partial Domains

The exogeneity theorem just proven concerns index numbers which satisfy certain properties on an unrestricted domain of data states. One might argue that it is only necessary, or even desirable, for these properties to be satisfied on a partial domain of data states. The exogeneity theorem can readily be extended to address this case. A minimal number of adjustments are required to the original formulation and proof. We will sketch the essential elements here.

Definition 25 *The general composition operator,* $\gamma : \mathbf{D}^n \longrightarrow \mathbf{D}$ *is a mapping which composes an ordered list of* n *original data states,* $\{d^{(j)}\}_{j\in\mathbf{N}}$*, to create a new data state, as follows:*

$$\gamma(\{d^{(j)}\}_{j\in\mathbf{N}}) = \{d' | d'_i = d_i^{(i)} \text{ for each } i \in \mathbf{N}\}. \tag{94}$$

Note that the original data states need not all be distinct.

Definition 26 *A subset of the set of data states,* $\mathbf{D}^* \subset \mathbf{D}$ *is closed under composition if:*

$$\text{If } d^{(i)} \in \mathbf{D}^* \text{for each } i \in \mathbf{N} \text{ then } \gamma(\{d^{(i)}\}_{i\in\mathbf{N}}) \in \mathbf{D}^* \tag{95}$$

Definition 27 $\mathbf{D}^* \subset \mathbf{D}$ *contains a continuum of prices if* $\exists\ p_{\min}, p_{\max}$ *such that*

$$\forall i \in \mathbf{N}, \forall p \in [p_{\min}, p_{\max}\ ]\ \exists d \in \mathbf{D}^* \text{ such that } d_{p_i} = p \tag{96}$$

Definition 28 $\mathbf{D}^* \subset \mathbf{D}$ *represents a partial domain of data states if* $\mathbf{D}^*$ *is closed under composition and* $\mathbf{D}^*$ *contains a continuum of prices.*

We can now reformulate all our definitions in terms of this partial domain $\mathbf{D}^*$:

An index number system S is anonymous, transitive, BCF, GNR or exogenous on $\mathbf{D}^*$ if these properties hold for the restricted set of data states $d \in \mathbf{D}^*$.

We can also restate the exogeneity theorem in stronger terms:

Theorem 4 THE PARTIAL DOMAIN EXOGENEITY THEOREM

If $\mathbf{D}^* \subset \mathbf{D}$ *is a partial domain of data states then* S *is exogenous for a convex evaluative standard on* $\mathbf{D}^*$ *if and only if* S *is anonymous, transitive, BCF and GNR on* $\mathbf{D}^*$*.*

Proof. The proof is a straightforward extension of the original 7.4. Since D is closed under composition, permutation and transcription operators can be applied without leaving D. Since D contains a continuum of prices, a weakly concave cost function can still be identified, which is defined on this subset of prices. The Legendre transform of this weakly concave function defined on a restricted domain is dual to a weakly convex evaluative standard on a dual restricted domain.[21]. This unique evaluative standard rationalizes all the index numbers that it is possible to generate over D, by virtue of transitivity. The details of the proof are unchanged beyond these straightforward modifications. ■ ■

8 Conclusions and Interpretation

The exogeneity theorem establishes that certain discrete properties are possessed by an index number system if and only if it can be interpreted as referring to a fixed evaluative standard, in which case we refer to it as exogenous. These discrete properties are bilateral comparison focus, anonymity, transitivity and global normative rationalizability. Bilateral comparison focus and anonymity are requirements concerning the relevance of the index numbers produced for the purpose at hand, the relevance of which must be judged in light of that purpose. Transitivity is a requirement concerning the internal consistency of the index numbers produced, which may be appropriate to impose under

[21] See Borwein and Lewis (2000), op cit.

certain conditions. Global normative rationalizability is a requirement concerning the meaningfulness of the index numbers to which the index number system gives rise. Although it is evident that an index number system which refers to a fixed evaluative standard has these discrete properties, it is far from obvious that the obverse is true.

The exogeneity theorem has significant conceptual and practical implications. It suggests that normative judgments are inescapable in the construction and use of index numbers. Index number systems which satisfy the discrete criteria identified above are exogenous and therefore produce index numbers which correspond to a single implicit or explicit normative criterion. Index number systems which fail to satisfy the discrete criteria identified above are either not globally normatively rationalizable, in which case at least some sets of index numbers that they generate fail to possess a meaningful welfare interpretation, or they correspond to necessarily different evaluative standards in different contexts (data states to which they may be applied), in which case comparisons across contexts are not always feasible without the use of a "meta-evaluative criterion" (identifying level sets of distinct evaluative relations as corresponding to an equivalent level of achievement) which again entails normative judgment.

The theorem is related to a deep-seated principle, which may be called the “comparison principle”. The comparison principle states that every statement of comparison involves some invariant concept. This is an observation about the very idea of comparison. When faced with a comparison, we may therefore reasonably ask, what is the concept which is being held invariant? Every index number system compares units of comparison to one another according to some invariant concept, if only the concept specified by the formal relation (defined by the index number system) in which each object of comparison stands to others. For example, the Fisher index number system compares objects of comparison to one another according to the "Fisher relation" in which they stand to one another, where the “Fisher relation” is defined as the outcome of the procedure of calculating the weighted average of prices using the quantity weights of each object of comparison in turn and then using the geometric mean as a method of averaging these two weighted averages. This purely formal type of invariance may not however be satisfactory from the standpoint of meaningfulness. We may seek some more substantive form of invariance, in order to be contented that the comparisons which we undertake are meaningful and relevant to the purpose at hand. Exogeneity is one such more substantive form of invariance. Similarly, a meta-evaluative relation which places distinct evaluative standards in some relation to one another is a more general substantive form of invariance (namely an invariant level of achievement of a specific meta-evaluative relation). The exogeneity theorem establishes a connection between these substantive forms of invariance and other desirable properties of an index number system.

One discrete property of an index number system (anonymity) demands that the index numbers generated by an index number system should depend on the data that it is applied to, and be invariant to the specific identities, or names, of the objects of comparison. A second discrete property of an index number system (transitivity) demands that the comparison between two objects of comparison should be invariant to whether it is undertaken directly or 'indirectly' by composing bilateral comparisons involving intermediate objects of comparison. A third discrete property of an index number system (bilateral comparison focus) demands that the index number generated for a specific pair of objects of comparison at a given "data state" should be invariant to changes in the data associated with other objects of comparison. A fourth discrete property of an index number system (global normative rationalizability) demands that an invariant evaluative standard may be used to rationalize the distinct index numbers generated by the index number system at a specific data state. The exogeneity theorem establishes a relation between these four invariances and another invariance (that of the evaluative criterion that may be used to interpret the index numbers produced by the index number system at different data states).

A central idea which is developed and used in the proof of the exogeneity theorem is the 'generalized within bounds' requirement. The generalized within bounds requirement states that a set of index numbers must be such that hypothetical agents associated with each object of comparison who are assigned incomes in proportion to the index numbers must not be assigned so much or so little income that one agent can buy strictly more of every good than the other agent. This 'non-domination' requirement can be derived straightforwardly from the idea that a set of index numbers must have an interpretation as generating an invariant level of achievement of some evaluative relation. If the non-domination requirement is violated, then a single evaluative relation satisfying minimal properties (such as that greater consumption is deemed to confer greater welfare) cannot be applied to all agents and still lead to the conclusion that they experience an equal level of achievement of this evaluative relation.

Many widely used index numbers fail to satisfy one of the discrete properties identified in the theorem, and therefore are either not globally rationalizable (i.e. they do not produce index numbers which satisfy the non-domination requirement in at least some specific data states) or are not exogenous (in which case the evaluative standard used to offer the index numbers a welfare interpretation may have to be changed as the data state changes). For example, the index number system which is defined by the ratio of nominal incomes possesses the properties of anonymity, bilateral comparison focus and transitivity, but is not globally normatively rationalizable. The index numbers it creates can violate the non-domination requirement. In contrast, well-known index number systems that possess the property of 'exactness' (which has been much discussed in the literature) are anonymous and transitive but violate the requirement of bilateral comparison focus and therefore are not exogenous. Whether an index number system satisfies bilateral comparison focus can be determined directly from its functional form. Whether an index number system is exogenous can also be determined directly from its functional form, as that form is necessarily "quotient separable" (i.e. of the form $\frac{c(p_1)}{c(p_2)}$) as shown above. This fact suggests a simple way to examine index number systems to determine whether they possess desirable properties and ought to be used. Widely used 'exact' index number systems also often fail to produce index numbers which satisfy the generalized within bounds requirement.[22] As a result, it may be impossible not only to provide a fixed welfare interpretation for the index numbers they produce[23] but it may be impossible (at specific data states) to provide for them any welfare interpretation at all. This fact provides a serious indictment of existing index number systems including those which have been celebrated as "exact".

There is an interesting and important relation between the discrete properties identified in the theorem and the properties of the "continuous" Divisia index. It is well-known that the Divisia index is path independent if and only if it is calculated along a path consisting of price-quantity combinations which reflect optimization of an exogenous "aggregation function", which may be interpreted as representing an evaluative relation[24]. The proof of this fact is contained in Appendix three. The result is a consequence of the well-known "potential function theorem". We have added to it by establishing the equivalence between the index number system's possession of a set of potentially desirable discrete properties, its possessing a fixed welfare interpretation, and the path independence of the Divisia index. The continuous Divisia index is equivalent to a discrete index under these conditions. Indeed, the exogeneity theorem therefore establishes a three-way equivalence between desirable properties of continuous index number systems, desirable properties

[22] See the accompanying paper, *The Choice of Index Number: Part II.*

[23] The concept of fixity here is that of a fixed utility function as opposed to a fixed family of utility functions. Exact index number systems may produce index numbers that can (where they are rationalizable at all) always be rationalized by *some* member of a specific family of utility functions, without it being the case that there is a *single* member of this family which rationalizes the index numbers that the index number system produces at all data states.

[24] See Hulten, Charles R. (1973), "Divisia Index Numbers," *Econometrica*, vol. 41(6), pp. 1017-25.

of discrete index number systems and the existence of a common fixed welfare interpretation.

Finally, it is interesting to note that the definition of global normative rationalizability does not demand that the evaluative standard used to rationalize the index numbers at each data state be convex. Rather it merely demands that a monotonicity requirement should not be violated at any given state. Nevertheless, the theorem ensures that, if the conditions of the theorem are satisfied, the resulting exogenous evaluative standard is convex. This may seem surprising. However, it is not in light of the fact that the convex evaluative standard is the inner envelope of all of the possible evaluative standards (convex and non-convex) that can rationalize the index numbers generated at a specific data state. Hence, under the specified conditions it is the *only* evaluative standard that can rationalize the index numbers generates at *all* possible data states.

We have shown that the index numbers generated by index number systems must satisfy the generalized within bounds requirement if these index numbers are to be rationalizable. However it can be shown (as is done in part II of this paper) that in practice most index number systems violate this requirement. This is a serious challenge to the meaningfulness and validity of most index number systems used in practice and suggests the need for deeper reflections about the foundations of applied economic work involving the use of index numbers.

The exogeneity theorem establishes a deep and unavoidable connection between the existence of a fixed evaluative standard which may be used to rationalize the index numbers generated by an index number system, and the possession by the index number system of potentially desirable properties. This result has wide-ranging implications for how we understand the relation between fact and value in economics, and for how we go about our practical attempts to describe the empirical world. The exogeneity theorem directs us to a basic truth: the task of valuation and the activity of evaluation are deeply and inescapably related.

9 Appendices

9.1 Intransitivity of Bilateral Index Number Systems: An Example

Suppose that there are three countries in the world - x, y, and z - and two commodities. Suppose that we wish to make comparisons of the aggregate price level across countries. In particular, we wish to arrive at index numbers representing, for example, the relative "cost of living" between country x and country y or the relative "cost of living" between country x and country z. Suppose that the following price structures and consumption patterns were observed (x currency is xen, y currency is yen, z currency is zen and everything is denominated in local currency units):

$$\overline{p}_x = (2,2), \overline{q}_x = (2,2) \tag{97}$$

$$\overline{p}_y = (1,4), \overline{q}_y = (3,1) \tag{98}$$

$$\overline{p}_z = (3,1), \overline{q}_z = (1,4) \tag{99}$$

How should the relative cost of living between countries x and y be identified? In particular, how many xen are required to confer the same "amount" of purchasing power as one yen? Putting the problem equivalently, how should we "scalarize" the price vectors we have observed?

One option is to use the bilateral Paasche index to compute the PPP between countries x and y:

$$S^P_{xy} = \frac{\langle \overline{p}_x, \overline{q}_x \rangle}{\langle \overline{p}_y, \overline{q}_x \rangle} = \frac{\langle (2,2),(2,2) \rangle}{\langle (1,4),(2,2) \rangle} = \frac{8}{10} = \frac{4}{5} = 0.800 \tag{100}$$

The inner product is a common way to scalarize vectors.[25] Why not use country x's own consumption pattern as scalarizing weights, as does the Paasche index? In this case, the index number conveys the relative cost in the two countries of purchasing country x's consumption bundle. Four xen and five yen are deemed to have the same purchasing power, according to this conception. However, is this the right approach to adopt? We could alternatively employ the bilateral Laspeyres index, which uses country y's consumption pattern as scalarizing weights:

$$S_{xy}^{L} = \frac{\langle \overline{p}_x, \overline{q}_y \rangle}{\langle \overline{p}_y, \overline{q}_y \rangle} = \frac{\langle (2,2),(3,1) \rangle}{\langle (1,4),(3,1) \rangle} = \frac{8}{7} = 1.143 \tag{101}$$

Now eight xen and seven yen are deemed to have the same purchasing power, so one xen has greater purchasing power than one yen - the opposite of the situation according to Paasche. Which approach should we prefer? Fisher famously suggested that we take the geometric mean of the Paasche and Laspeyres, thus "deferring" to the respective countries in a diplomatic manner:

$$S_{xy}^{F} = \sqrt[2]{S_{xy}^{p} * S_{xy}^{L}} = \sqrt[2]{\frac{8}{10} * \frac{8}{7}} = 0.956 \tag{102}$$

Under this approach to comparison, one xen and one yen have just about the same purchasing power. We can also use the Fisher index to calculate the relative cost of living between countries x and z, and between countries y and z:

$$S_{xz}^{F} = \sqrt[2]{\frac{\langle \overline{p}_x, \overline{q}_x \rangle}{\langle \overline{p}_z, \overline{q}_x \rangle} * \frac{\langle \overline{p}_x, \overline{q}_z \rangle}{\langle \overline{p}_z, \overline{q}_z \rangle}} = \sqrt[2]{\frac{8}{8} * \frac{10}{7}} = 1.195 \tag{103}$$

$$S_{yz}^{F} = \sqrt[2]{\frac{\langle \overline{p}_y, \overline{q}_y \rangle}{\langle \overline{p}_z, \overline{q}_y \rangle} * \frac{\langle \overline{p}_y, \overline{q}_z \rangle}{\langle \overline{p}_z, \overline{q}_z \rangle}} = \sqrt[2]{\frac{7}{10} * \frac{17}{7}} = 1.304 \tag{104}$$

[25] Indeed, any linear operator that does so can be represented as an inner product, according to the well-known Riesz Representation Theorem.

Thus one needs about 1.2 xen or 1.3 yen to enjoy the same purchasing power as that conferred by one zen. However, note that:

$$\frac{S^F_{xz}}{S^F_{yz}} = \frac{1.195}{1.304} = 0.916 \neq 0.956 = S^F_{xy} \tag{105}$$

The direct comparison between countries x and y yields a different answer than does the indirect comparison via country z; the Fisher index is intransitive.

This problem can be easily overcome, by employing the EKS multilateral price index, which takes the geometric mean of Fisher bilateral comparisons:

$$S^{EKS}_{xy} = \sqrt[3]{(S^F_{xy})^2 * \frac{S^F_{xz}}{S^F_{zy}}} = \sqrt[3]{(0.956)^2 * \frac{1.195}{1.304}} = 0.943 \tag{106}$$

If we computed the indirect comparison between countries x and y via country z using the EKS multilateral price index, we would get exactly the same answer as in the direct comparison: 0.943. The EKS index is *designed* to be transitive. The EKS index number system is one of many employed in practice. It possesses a number of properties, such as transitivity, but lacks other desirable properties. It can be shown that the index numbers produced by the EKS index number system are consistent with a specific welfare interpretation under restricted conditions[26] and that whenever such a welfare interpretation exists, the Fisher index is itself transitive, making recourse to the EKS index unnecessary[27]. It is less well known, but also true that these index numbers are inconsistent with *any* reasonable welfare interpretation under other conditions. As we demonstrate in the paper, this fact is simply an instance of a general phenomenon.

[26] As shown in e.g. Diewert (1996), "Axiomatic and Economic Approaches to International Comparisons", NBER Working Paper number 5559.

[27] See e.g. Neary, J. Peter (2004), "Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income", *American Economic Review*, Vol. 94 No. 5.

9.2 Transitivity and Related Properties

In this appendix, we lay out the relationship between transitivity and related potentially desirable properties. For example, we may require that the relative cost of living between a given object of comparison and itself should be identically equal to unity:

Definition 29 *S is self-consistent if the following equality holds:*

$$S_{xx}(d) = 1 \qquad x \in \mathbf{N}, \forall d \in \mathbf{D} \tag{107}$$

Note that this property of the index number system imposes a restriction on the index number set produced by the index number system at every data state. Properties of index number sets can similarly be extended to those of index number systems.

Another property in which we may be interested is the following:

Definition 30 *S is role-invariant if the following equality holds:*

$$S_{xy}(d) = \frac{1}{Syx(d)} \qquad \forall x, y \in \mathbf{N}, \forall d \in \mathbf{D} \tag{108}$$

This property says something more substantive; if we switch the role of focal and reference object of comparison the new index number should be the multiplicative inverse of the old one. We take the multiplicative inverse because the index number is meant to express the ratio of the two costs of living. Clearly, role-invariance subsumes self-consistency:

Lemma 12 *If S is role-invariant then S is self-consistent.*

Proof. Let $x = y$ in 40 and note that S_{xx} is a positive number. ■ ■

Both of these properties seem eminently desirable. It is tempting to consider them minimal properties that any reasonable index number system should have. However, there is a significant difference between self-consistency and role-invariance. The former property stems from the fact that we are comparing two situations; we demand that our method of comparison recognize when two situations are identical. Given that our index numbers are meant to convey a multiplicative comparison, and that unity is the multiplicative identity, it is hard to argue with self-consistency. This is not the case with role-invariance. Role-invariance should hold if focal and reference roles merely indicate which object of comparison should be in the numerator and which object of comparison should be in the denominator of the multiplicative comparison and if the relevant features of the object being compared are the same in both instances. However, what if the position of the objects of comparison also indicated the perspective adopted by the evaluator when making the comparison (and in particular the feature of the object being compared)? Perhaps the relative cost of living between the China and India is 2 from the Indian "perspective" and the relative cost of living between India and the China is 0.4 from the Chinese "perspective". This "multi-perspectival" approach is perfectly intelligible, but it violates role-invariance, because role "matters".

Clearly, it is simpler to adopt one perspective than to adopt many. The goal of constructing index numbers is after all to reduce the multiplicity of dimensions of information with which we are faced to a single dimension of information in order that unambiguous statements can be made. We much prefer newspaper headlines that read "Growth higher in China than in India" to ones that reads "Growth Higher in China than in India (According to the Chinese) and Growth Higher in India than in China (According to the Indians)". The preference for unambiguous characterizations of the world is understandable, if not always ultimately justifiable.

The critical point for the present purpose is that self-consistency requires that the index number express a multiplicative comparison of price situations whereas role-invariance requires that the perspective of comparison not vary with the role each object of comparison plays in the comparison. A more substantive restriction would extend this invariance to the role that a third object of comparison might play in indirect comparison:

Definition 31 *S is base invariant if the following equality holds:*

$$\frac{Sxw(d)}{Syw(d)} = \frac{Sxz(d)}{Syz(d)} \qquad \forall w, x, y, z \in \mathbf{N}, \forall d \in \mathbf{D} \tag{109}$$

A related requirement is that the indirect comparison give the same answer as the direct comparison.

Definition 32 *S is transitive if the following equality holds:*

$$S_{xy}(d) = \frac{Sxz(d)}{Syz(d)} \qquad \forall x, y, z \in \mathbf{N}, \forall d \in \mathbf{D} \tag{110}$$

Clearly, these various properties are intimately related:

Lemma 13 *If S is self-consistent and base invariant, then I is transitive.*

Proof. Let $w = y$ in 29 and apply 27: ■

$$S_{xy}(d) = \frac{Sxy(d)}{1} = \frac{Sxy(d)}{Syy(d)} = \frac{Sxz(d)}{Syz(d)} \qquad \forall x, y, z \in \mathbf{N}, \forall d \in \mathbf{D} \tag{111}$$

This statement can in fact be extended to an "if and only if" form, as transitivity implies the other three properties:

Lemma 14 *If S is transitive then S is self-consistent, role-invariant and base invariant.*

Proof. Suppose S is transitive. Then, by 30: ■

$$\frac{Sxw(d)}{Syw(d)} = S_{xy}(d) = \frac{Sxz(d)}{Syz(d)} \qquad \forall w, x, y, z \in \mathbf{N}, \forall d \in \mathbf{D} \tag{112}$$

Thus S is base invariant. Now, let $x = y = z$ in 30:

$$S_{xx}(d) = \frac{Sxx(d)}{Sxx(d)} = 1 \qquad \forall x \in \mathbf{N}, \forall d \in \mathbf{D} \tag{113}$$

Thus S is self-consistent. Now, let $x = z$ in 30 and apply 27:

$$S_{xy}(d) = \frac{Sxx(d)}{Syx(d)} = \frac{1}{Syx(d)} \qquad \forall x, y \in \mathbf{N}, \forall d \in \mathbf{D} \tag{114}$$

Thus S is role-invariant. ■

Thus, transitivity subsumes the other three properties. This discussion may clarify why transitivity is a standard property of index number systems. Its conceptual import is that the index numbers produced should convey a bilateral multiplicative comparison, the content of which depends solely on the identity of the two objects of comparison being compared, not on their respective roles in the comparison, nor the identity of any third object of comparison via which a bilateral comparison might "indirectly" be made.

9.3 Path-Independence of the Divisia Index and the Existence of a Fixed Welfare Intepretation

First, note that the following conditions are equivalent, and form the definition of a conservative vector field, $\widehat{F}$:

1. For any oriented simple closed curve C, the line integral $\oint_C \widehat{F} \cdot ds = 0$

2. For any two oriented simple curves with the same endpoints, the line integrals over these two curves are the same, i.e. $\int_{C_1} \widehat{F} \cdot ds = \int_{C_2} \widehat{F} \cdot ds$

3. There exists a scalar "potential function" $\widehat{f}$ such that $\widehat{F} = \bigtriangledown \widehat{f}$ where $\bigtriangledown$ is the gradient.[28]

We refer to this equivalence as the "potential function theorem", although it is sometimes referred to by other terms such as the "gradient theorem".

Second, consider the function from the vector space of prices $\mathbf{P}$ to the set of real numbers which is defined by $c(p, E = \overline{E})$, which refers to the minimum cost of achieving satisfaction level $\overline{E}$ given evaluative function E (assumed to be "well-behaved" in the sense otherwise presupposed in our analyses, i.e. it possesses commodity bundles as arguments is monotonic etc.) and prices p. For ease of notation, we refer to this minimum cost function henceforth simply as $c(p)$, keeping the evaluation function and achievement level in the background. Now, consider the ("potential") function defined by $\Phi(p) = \ln c(p)$ and the associated vector field defined by $\widehat{F}_{x^*} = \bigtriangledown \ln c(p) = \left(\frac{1}{c(p)}\frac{\partial c}{\partial p_1}, ..\frac{1}{c(p)}\frac{\partial c}{\partial p_n}\right)$. This vector field has a number of interesting features. Note that by Shephard's lemma, the quantities that are optimal to consume given a price vector, p, are given by $x^*(p) = (x_1^*(p), ..x_n^*(p)) = \left(\frac{\partial c}{\partial p_1}, ..\frac{\partial c}{\partial p_n}\right)$, where $x^*(p)$ is a bundle of commodities, in a commodity space X. Therefore, the vector $\left(\frac{1}{c(p)}\frac{\partial c}{\partial p_1}, ..\frac{1}{c(p)}\frac{\partial c}{\partial p_n}\right)$ represents the amounts of each commodity being consumed per unit of expenditure.

Now, we could define an index number system as follows: $I_{jk} = e^{\Phi(p_j) - \Phi(p_k)} = \frac{c(p_j)}{c(p_k)}$. Now, by the potential function theorem, for any two oriented simple curves with the same endpoints, the line integrals over these two curves are the same, i.e. $\int_{C_1} \widehat{F}_{x^*} \cdot ds = \int_{C_2} \widehat{F}_{x^*} \cdot ds$, where the curves C are within $\widehat{F}_{x^*}$. By definition of the line integral, $\int_C \widehat{F}_{x^*} \cdot ds = \int_C \sum_{i=1}^n \left(\frac{1}{c(p)}\frac{\partial c}{\partial p_i}\right) dp_i$. If the quantities consumed at each price are optimal given the evaluation function, then Shephard's

[28] (see, for example, http://mathworld.wolfram.com/ConservativeField.html or other texts).

lemma holds and we can write: $\int_C \widehat{F}_{x^*} \cdot ds = \int_C \sum_{i=1}^n \frac{x_i^*(p)}{c(p)} dp_i$. Finally, if C ranges from point p_k to point p_j then, by the relationship between the line integral and the potential function specified by the potential function theorem, $\ln I_{jk} = i(p_j) - i(p_k) = \int_C \sum_{i=1}^n \frac{x_i^*(p)}{c(p)} dp_i$ and this is independent of the specific curve C as long as it runs between the two endpoints concerned. In order to interpret the expression, let us rewrite it with one further step so that it reads: $\ln I_{jk} = \Phi(p_j) - \Phi(p_k) = \int_C \sum_{i=1}^n \frac{x_i^*(p)p_i}{c(p)} \frac{dp_i}{p_i} = \int_C \sum_{i=1}^n \theta_i(p) \frac{dp_i}{p_i}$

where $\theta_i(p) = \frac{x_i^*(p)p_i}{c(p)}$. This tells us that the log of the "index number" between the two endpoints in the space of prices is the sum (along the path between these two points) of the weighted average "growth" in prices of commodities, where the growth in prices of individual commodities is weighted by the share in total expenditure of the expenditure on each commodity at each point along the path.

One further step gives us $I_{jk} = e^{\int_C \sum_{i=1}^n \theta_i(p) \frac{dp_i}{p_i}}$. Now, this is nothing more than the classical Divisia price index.[29]

So we have shown how the Divisia price index arises from the potential function interpretation of index numbers and the assumption that the quantities to which it is applied are the optimal ones given the prices that are experienced and a fixed evaluation function (in relation to which the potential function is defined). Indeed, by the potential function theorem, a Divisia price index has a value independent of the path between two end points if and only if there exists an evaluation function (what he calls an aggregation function) such that the quantities and prices are related in exactly this way.[30]

Now, there is one more interesting interpretation of this result. When there exists a potential function interpretation of the index numbers, then $\ln I_{ik} = \Phi(p_i) - \Phi(p_k)$ is independent of the "path" between i and k. In particular, $\ln I_{ik} = \Phi(i) - \Phi(j) + \Phi(j) - \Phi(k) = \ln I_{ij} + \ln I_{jk}$ which is equivalent to the statement that $I_{ik} = I_{ij} I_{jk}$ (i.e. transitivity of index numbers). Base country invariance and certain other standard axioms can be similarly derived. The contribution of the paper is not to show that when there exists a potential function interpretation of index numbers that then various such "discrete" properties of index numbers hold, but rather to show that if a

[29] [To verify this, see for example eq(1) in Trivedi (1981) in *International Economic Review*, Vol. 22, No. 1 (Feb., 1981), pp. 71-77. We can also arrive at the same observation by inverting prices and quantities in equation (1) in Hulten (1973), *op cit*].

[30] Hulten(1973) adds the requirement of linear homogeneity of the evaluation function, because he is dealing with production rather than price indices and therefore requires a means of cardinally comparing the extent of achievement embodied in different indifference surfaces.

certain set of "discrete" properties hold, then there exists a potential function interpretation of index numbers.

10 Figures

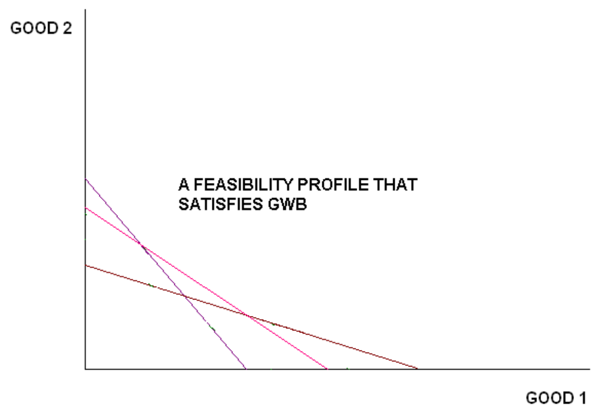

Figure 1:

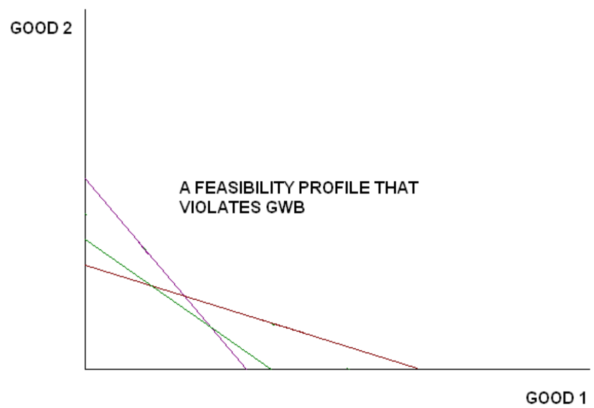

Figure 2:

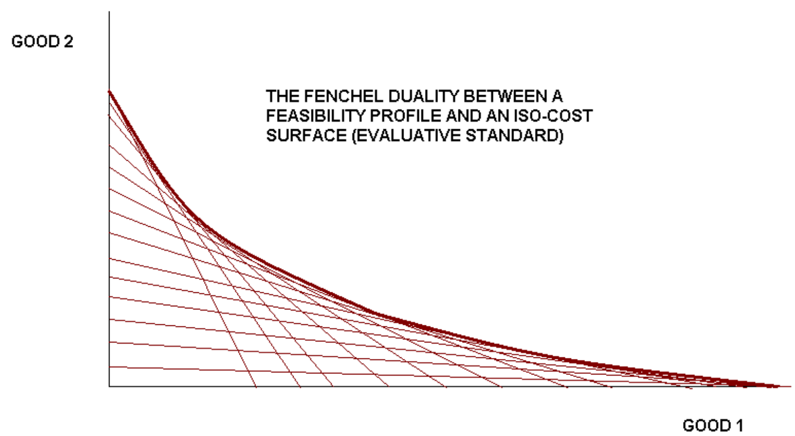
GOOD 2
THE FENCHEL DUALITY BETWEEN A
FEASIBILITY PROFILE AND AN ISO-COST
SURFACE (EVALUATIVE STANDARD)
GOOD 1

Figure 3: