CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 949

September 2009

How to Measure Living Standards and Productivity

Nicholas Oulton

Abstract

This paper sets out a general algorithm for calculating true cost-of-living indices or true producer price indices when demand is not homothetic, i.e. when not all expenditure elasticities are equal to one. In principle, economic theory tells us how we should calculate a true cost-of-living index or Konüs price index: first estimate the consumer's expenditure function (cost function) econometrically and then calculate the Konüs price index directly from that. Unfortunately this is impossible in practice since real life consumer (producer) price indices contain hundreds of components, which means that there are many more parameters than observations. Index number theory has solved this problem, at least when demand is homothetic (all income elasticities equal to one). Superlative index numbers are second order approximations to any acceptable expenditure (cost) function. These index numbers require data only on prices and quantities over the time period or cross section under study. Unfortunately, there is overwhelming evidence that consumer demand is *not* homothetic (Engel's Law). The purpose of the present paper is to set out a general algorithm for the nonhomothetic case. The solution is to construct a chain index number using *compensated*, not actual, expenditure shares as weights. The compensated shares are the actual shares, adjusted for changes in real income. These adjustments are made via an econometric model, where only the responses of demand to income changes need to be estimated, not the responses to price changes. This makes the algorithm perfectly feasible in practice. The new algorithm can be applied (a) in time series, e.g. measuring changes over time in the cost of living; (b) in cross section, e.g. measuring differences in the cost of living and hence the standard of living across countries; and (c) to cost functions, which enables better measures of technical progress to be developed.

Keywords: consumer price index, Konüs, cost of living, measurement of welfare change, Quadratic Almost Ideal Demand System, producer price index, homothetic, productivity
JEL Classifications: C43, D11, D12, E31, D24, I31, O47

This paper was produced as part of the Centre's Productivity and Innovation Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements

I owe thanks to Erwin Diewert for detailed comments and helpful suggestions; he is not responsible for my conclusions or any errors. I am grateful also to the U.K. Economic and Social Research Council who have financed this research under ESRC grant number RES- 000-22-3438.

Nicholas Oulton is an Associate at the Centre for Economic Performance, London School of Economics. He is also Visiting Professor in the Department of Economics, University College London.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

ISBN 978-0-85328-415-4

1. Introduction

This paper sets out an algorithm for measuring the true cost of living in the important case where demand is non-homothetic. The algorithm can be applied both to time series and to cross sections, eg cross-country studies of living standards. Essentially the same algorithm can be applied to the parallel problem of measuring the price of producers' inputs, which in turn is a step on the road to measuring technical progress. The algorithm is practical since it requires no more data than is needed to calculate conventional index numbers. And in principle it can be implemented at the same level of product detail at which conventional index numbers are constructed by national statistical agencies.

Economic theory tells us how to measure the true cost of living: estimate the expenditure function econometrically and then calculate the Konüs price index. The Konüs price index for period t relative to some other period r is defined as the ratio of the (minimum) cost of achieving a given utility level at the prices of period t to the cost of achieving the same utility level at the prices of period r (Konüs, 1939). If we know the expenditure function then we can calculate the Konüs price index, for any chosen utility level. Similarly, economic theory tells us how to measure the true index of the cost of a producer's inputs: estimate the producer's cost function and calculate the analogue of the Konüs price index. If we know the cost function then we also know the degree of economies of scale, the size of any input biases in economies of scale, the growth rate of technical progress, and the size of any input biases in technical progress.

Though much work has been done on estimating systems of consumer demand or producers' cost functions, the results of these studies are not typically employed by other economists in empirical work. For example, when macro economists study inflation empirically, they do not usually employ their micro colleagues' estimates of expenditure functions. Rather they use consumer price indices constructed by national statistical agencies. The reason is clear. The economic approach cannot be applied at a level useful for other empirical economists because of data limitations.

1.1 The data problem

The economic approach cannot be employed because the number of parameters to be estimated is large and the number of observations is comparatively small. In other words the problem is a purely practical one which might in theory be solved just by waiting long enough (possibly for hundreds of years). This causes a dilemma for the empirical economist who is unwilling to wait. Either the economic approach must be abandoned and index numbers employed instead. Or the data must be aggregated and the economic approach applied at a higher level. The first way, I shall argue later, is perfectly all right if demand (for consumer goods or producer inputs) is homothetic. But if it is not, then index numbers will not measure what they are supposed to measure. The second approach is more relevant to testing economic theory rather than to using it. In practice, empirical economists tend to use the index numbers (for output, inputs and prices) supplied to them by statistical agencies, without asking too many questions about the assumptions on which they are based.[1]

The data problem can be illustrated by taking the Quadratic Almost Ideal Demand System (QAIDS) for N products of Banks, Blundell and Lewbel (1997) as an example. In the expenditure function of this system there are $\frac{1}{2}(N-1)(N+2)$ independent parameters relating to the consumer's response to prices and $2(N-1)$ independent parameters relating to the consumer's response to income, for a total (excluding a scale parameter) of $\frac{1}{2}(N-1)(N+6)$ independent parameters. The QAIDS is a system of $N-1$ independent equations for the expenditure shares. Roughly speaking, each of these equations contains on average $\frac{1}{2}(N+2)$ independent coefficients relating to prices and two coefficients relating to income. To have any chance of estimating these coefficients econometrically we must have more observations than coefficients; ie if we have T aggregate time series observations, then we require $T > \frac{1}{2}(N+6)$.

This is where the empirical study of demand and the practice of index number construction part company. National statistical agencies construct their indices of the cost of living from hundreds of components. For example, the U.S. Bureau of Labor Statistics constructs its Consumer Price Index from 305 "entry-level items" (U.S. Bureau of Labor Statistics, 2007). The U.K.'s Consumer Prices Index and Retail Prices Index have some 650 "items" (Office for National Statistics, 1998 and 2006). To estimate the parameters of the QAIDS for 650 products would require over three centuries of annual data, a requirement that

[1] See for example the remarks of Tobin (1987) on the contributions of Irving Fisher to index number theory: "These index number issues do not seem as important to present-day economists as they did to Fisher. Knowing that they are intrinsically unsolvable, we finesse them and use uncritically the indexes that government statisticians provide". Of course, I do not agree that these "index number issues" are "intrinsically unsolvable", otherwise I would not have written this paper.

is not and is never likely to be met. So when econometricians use time series data to test the theory of demand, they are forced to aggregate the products into a small number of groups; eg Christensen *et al*. (1975) tested the theory of demand using three product groups over 1929-72. But additional, strong assumptions are needed to justify this aggregation and these assumptions cannot be tested directly (Deaton and Muellbauer, 1980b, chapter 5). So the "prices" and "quantities" which are the basic data for testing the theory of demand in this kind of study are themselves index numbers.[2] But then the theoretical justification for these index numbers is unclear. Cross section studies of household demand fare better since in any given year it may be reasonable to assume prices are the same for all households (except for regional effects). With typically several thousand observations in any cross section, lack of observations is not a problem. But then only the effects of income (and of household composition) on demand can be measured, as in eg Blow, Leicester and Oldfield (2004).

The upshot is that all the empirical work that economists have done on household demand has had no effect on the measurements actually made by national statistical agencies (although the underlying theory may have been influential). Similar remarks apply to the measurement of other indices such as the producer price index.

1.2 Non-homotheticity

Actually, none of this matters much *provided that demand (for inputs or consumer goods) is homothetic*. If this condition holds and if we are prepared to accept that economic theory is true,[3] then we have no need to estimate cost or expenditure functions. We can instead use the superlative index numbers of Diewert (1976). As discussed more fully in section 2, these provide second order approximations to any acceptable utility or cost function.

Unfortunately, an overwhelming body of empirical evidence establishes that consumer demand is *not* homothetic. The most obvious manifestation of this is Engel's Law: the proportion of total household expenditure devoted to food falls as expenditure rises. Since its original publication in 1857, Engel's Law has been repeatedly confirmed. Houthakker (1957)

---

[2] Cross section studies also often employ highly aggregated data: five product groups in the case of Banks *et al.* (1997), eight in the case of Blundell *et al*. (2007), both studies of British household budgets, and 11 in the case of Neary (2004), a cross-country study of PPPs. The panel study on Canadian households of Lewbel and Pendakur (2009) employed nine groups.

[3] Throughout this paper I adopt the economic approach to index numbers; see Diewert (1981) and (2008) for surveys of this and of the alternative axiomatic and stochastic approaches, also Balk (1995) on the axiomatic approach.

showed that the Law held in some 40 household surveys from about 30 countries.[4] Engel's Law also holds in the much more econometrically sophisticated study of Banks *et al*. (1997) on UK household budgets. The prevalence of non-homotheticity is also confirmed by the more disaggregated studies of Blow *et al.* (2004), also on U.K. household budgets, which considered 18 product groups, and Oulton (2008) who considered 70 product groups.[5]

If demand is not homothetic, then superlative index numbers are not guaranteed to be good approximations to Konüs price indices, even locally. In fact the true price index may lie outside the Paasche-Laspeyres spread. And the true price index is no longer unique but depends on the reference level chosen for utility (or, for the producer price index, on the reference output level). The fact that the Konüs price index generally varies with the reference utility level is sometimes taken as puzzlingly paradoxical. But it can be given a simple intuitive justification. Consider a household with a very low standard of living spending 60% of its budget on food (as was the case with the working class households studied by Engel in 1857). Suppose the price of food rises by 20%, with other prices constant. Then money income will probably have to rise by close to (0.60 x 20% =) 12%, to leave utility unchanged, since there are limited possibilities for substituting clothing and shelter for food. Compare this household to a modern day British one, spending 15% of its budget on food prepared and served at home (Blow *et al*., 2004). Now the maximum rise in income required to hold utility constant is only (0.15 x 20 =) 3% and probably a good bit less as substitution opportunities are greater.

This leaves the welfare interpretation of conventional consumer price indices and their cross-country cousins, the Purchasing Power Parities (PPPs) constructed by the OECD and the World Bank, somewhat up in the air. If the true price index depends on the reference level

---

[4] Engel's (1857) results for expenditure by households of various income levels in Saxony are described more accessibly in Marshall (1920), chapter IV. In each of the surveys that he collected Houthakker (1957) estimated the elasticity of expenditure on food and three other groups (clothing, housing and miscellaneous) with respect to total expenditure and to household size. For each product group, he regressed the log of expenditure on that group on the log of total expenditure and the log of family size. He used weighted least squares on grouped data; individual data was not available to him. The results for food were clear-cut: demand was inelastic with respect to expenditure in every survey. The results for clothing and miscellaneous were equally clear-cut: demand was expenditure-elastic. The result for housing was more mixed.

[5] An exception to this consensus is Dowrick and Quiggin (1997). They studied the 1980 and 1990 PPPs for 17 OECD countries, using 38 components of GDP, and argued that the data could be rationalised by a homothetic utility function. But their anomalous finding may be due partly to the fact that the per capita incomes of these countries were fairly similar and partly to the low power of their nonparametric test (Neary, 2004).

of utility, how are we to interpret real world price indices? The answer in the time series context is that a chained, superlative index is likely to be approximately equal to a true price index with reference utility level at the midpoint of the sample period (Diewert, 1976 and 1981; Feenstra and Reinsdorff, 2000; Balk 2004).[6] For a cross-country comparison, the viewpoint will be that of a "middle" country. While there is nothing wrong with this viewpoint, there is no special reason why the midpoint should be so privileged. There is also the disadvantage that when the sample period is extended (or the number of countries in the comparison increased), the viewpoint changes.

A parallel issue arises on the production side and takes the form of input biases in economies of scale: if output is doubled, holding prices and technology constant, does that leave all cost shares unchanged? The possibility that this is *not* the case has certainly been entertained as a matter of theory, though I am not aware of any substantial body of empirical work devoted to this issue. But such a situation may be quite common. Consider a firm which has fixed and variable costs, where the fixed costs are white collar workers and the variable costs are blue collar workers. Then an expansion of output will lower the share of white collar workers in total costs. In this case the cost function is non-homothetic and also non-homogeneous in output. So it would certainly seem desirable to take non-homotheticity into account when trying to measure TFP.

1.3 The algorithm

The proposed algorithm can be summarised as follows. The growth rate of a Konüs consumer price index resembles that of a Divisia index (or the latter's empirical counterpart, a chain index) in that it is an expenditure-share-weighted average of the growth rates of the

---

6 Suppose a utility function exists which rationalises the data but may be non-homothetic. Diewert (1981) showed that there exists a utility level which is intermediate between the levels at the endpoints of the interval under study such that a Konüs price index over this interval, with utility fixed at the intermediate level, is bounded below by the Paasche and above by the Laspeyres. Balk (2004) showed that when the growth of prices is piecewise log linear a chained Fisher price index approximates a Konüs price index over an interval when the reference utility level is fixed at that of some intermediate point in the interval. More precise results are available for specific functional forms. Diewert (1976) showed that a Törnqvist price index is exact for a nonhomothetic translog cost function when the reference utility level is the geometric mean of the utility levels at the endpoints; see also Diewert (2009) for extensions. For the AIDS, Feenstra and Reinsdorf (2000) showed that, if prices are growing at constant rates, the Divisia index between two time periods equals the Konüs price index when the reference utility level is a weighted average of utility levels along the path.

component prices. But for the Konüs index the shares are not the actual, observed ones, but rather what I call the compensated shares: the shares that would be observed if prices were the actual, observed ones but utility were held constant at some given reference level. I derive a relationship between the compensated and the actual shares: the compensated shares are equal to the actual ones, adjusted for the difference in real income (utility) between the actual situation and the reference level. The adjustment requires us to know, for each product, the consumer's response to real income changes but not the response to price changes. This is why the algorithm can be implemented at a very disaggregated level, since the number of parameters needed to describe the consumer's response to income changes is quite small: in the case of the QAIDS only two parameters for each product need to be known. These income response parameters can be estimated econometrically, provided we do not try at the same time to estimate the responses to individual price changes. This can be done by estimating a flexible demand system such as the QAIDS but with the price variables replaced by a much smaller number of principal components. In this way the data limitation problem can be overcome.

It is important to note that the algorithm proposed here is not designed as a test of whether the theory of consumer (or producer) demand is true. Rather it seeks to *use* demand theory to construct better measures of living standards and productivity. In fact, the algorithm assumes that demand theory is true and hence that the consumer's or producer's responses can be approximated by a flexible system like the QAIDS.

1.4 Plan of the paper

I start in section 2 by reviewing the theory of superlative index numbers. I argue that these solve the problem of estimating a true price index in the homothetic case. In section 3 I go on to consider the non-homothetic case and present a general algorithm for estimating a true (Konüs) price index for a representative consumer. The algorithm requires just the same data (and no more) as would be required to estimate a conventional index number. This algorithm is illustrated more specifically for the QAIDS. I argue that it can be applied both to time series and to cross section (eg cross country studies). In section 4 the analysis is extended by dropping the assumption of a representative consumer. I show how the QAIDS can be adapted to allow for inequality in the distribution of income. It turns out that this just requires adding two additional variables, both statistics of the income distribution, to the share equations of the QAIDS. The algorithm derived for the simpler case of a representative consumer can then be applied much as before. This section also discusses including household characteristics as additional determinants of demand. Section 5 shows how the general method applies, after some adaptation, to the estimation of a true input price index for producers, in the case where economies of scale may exist and may be input-biased. A true input price index is a step on the road to estimating the growth rate of technical progress, which may also be input-biased. The algorithm enables input biases in economies of scale and in technical progress to be estimated simultaneously. Finally, section 6 concludes.

2. Price indices: the homothetic case

In this section I argue that superlative index numbers have solved the problem of measuring the true cost of living for a single, representative consumer in the case where demand is homothetic.

Let the consumer's expenditure function be

$$x = E(\mathbf{p}, u)$$

This shows the minimum expenditure x needed to reach utility level u when $\mathbf{p} = (p_1\, p_2 \ldots p_N)$ is the Nx1 price vector faced by the consumer**;** $x = \sum_i p_i q_i$ where the q_i are the quantities purchased. Expenditure at time t is therefore a function of prices at time t and the utility level. Suppose that, hypothetically, utility were held at its level at time b while the consumer faced the prices of time t. Let $x(t,b)$ denote the minimum expenditure at the prices of time t required to achieve the utility level of time b. Then

$$x(t,b) = E(\mathbf{p}(t), u(b)) \tag{1}$$

For brevity write the right hand side as

$$E(t,b) = E(\mathbf{p}(t), u(b))$$

where the first argument of $E(t,b)$ is the time period for prices and the second is the time period for utility. The Konüs price index at time t, with time b as the base period for utility, is defined as the ratio of the minimum expenditure required with the prices of time t to attain the utility level of time b, to the minimum expenditure required to attain this same utility level, when the consumer faces the prices of time b:[7]

[7] It is convenient if the reference period for the Konüs price index (the period when the index equals 1) is the same as the base period. But nothing important would be changed if we

$$P^K(t,b) = E(t,b)/E(b,b) \quad (2)$$

(Clearly, $P^K(b,b) = 1$). In general, the Konüs price index depends on both the prices and the specified utility level. However as is well known, the index is independent of the utility level and depends only on the prices if and only if demand is homothetic, ie if all income elasticities are equal to one (Konüs, 1939; Samuelson and Swamy, 1974; Deaton and Muellbauer, chapter 7, 1980b).

I wish to argue that the problem of estimating true cost-of-living indices and indices of the standard of living, together with their counterparts on the production side, has been solved in the case where demand is homothetic, at least within the limit of what is empirically possible. The solution was in fact provided by Diewert's development of superlative index numbers, index numbers which are exact for some flexible functional form (Diewert, 1976). A flexible functional form is one which provides a second order approximation to any expenditure function (or utility function) or to any cost function (or production function) which is acceptable to economic theory.[8] Note that these are local not global properties; a good approximation at the point in question does not guarantee a good approximation at some other point.

The flexible functional forms which Diewert (1976) studied were what he called quadratic means of order s, given by:

$$A(\mathbf{p};s) = \left[\sum_{i=1}^{i=N}\sum_{j=1}^{j=N} b_{ij} p_i^{s/2} p_j^{s/2}\right]^{1/s}, \; b_{ij} = b_{ji}, \forall i \neq j, s > 0 \quad (3)$$

where $A(\mathbf{p};s)$ is assumed concave and positive. For concreteness, in this section I interpret equation (3) as referring to the consumer's problem of choosing amongst N products subject to a budget constraint but it could equally well refer to the producer's problem of allocating a given expenditure amongst N inputs. Under this interpretation, $A(\mathbf{p};s)$ is the cost per unit of utility and equation (3) is part of an expenditure function of the following form:

$$x(t,b) = A(\mathbf{p}(t);s)u(b) \quad (4)$$

---

chose the reference year to be year r and defined the Konüs price index with base period b and reference period r as $P^K(t,b,r) = E(t,b)/E(r,b) = [E(t,b)/E(b,b)]/[E(r,b)/E(b,b)] = P^K(t,b)/P^K(r,b)$.

8 A second order approximation is one for which the approximating function and the function approximated have the same value at a particular point, the first derivatives of the two functions are equal at that same point, and the second derivatives are also equal at that point.

where $x = \sum_i p_i q_i$ is total expenditure, q_i is the quantity purchased of the *i*th product, and $x(t,b)$ is the minimum expenditure required to reach the utility level prevailing at time *b* when the consumer faces the prices of time *t*. Note that equation (4) implies that demand is homothetic: all expenditure elasticities are equal to one.[9]

The Konüs price index for period *t* relative to period *b* corresponding to this expenditure function is then

$$P^K(t,b) = x(t,b) / x(b,b) = A(\mathbf{p}(t);s) / A(\mathbf{p}(b);s)$$

which is independent of the utility level. If the consumer maximises utility subject to the budget constraint $x(t,t) = \sum_i p_i(t) q_i(t)$, then Diewert showed that the Konüs price index for period *t* relative to period *b* which corresponds to (3) is given by:

$$P_s(t,b) = \left[\frac{\sum_{i=1}^{N} \left(\frac{p_i(t)}{p_i(b)} \right)^{s/2} \left(\frac{p_i(b) q_i(b)}{\sum_{i=1}^{i=N} p_i(b) q_i(b)} \right)}{\sum_{i=1}^{N} \left(\frac{p_i(b)}{p_i(t)} \right)^{s/2} \left(\frac{p_i(t) q_i(t)}{\sum_{i=1}^{i=N} p_i(t) q_i(t)} \right)} \right]^{1/s} \tag{5}$$

Note that base period (period *b*) expenditure shares appear in the numerator and current period (period *t*) ones in the denominator.

The importance of this result is that the formula for the price index requires knowledge only of prices and quantities (or equivalently, prices and budget shares). It does not require knowledge of any of the parameters of $A(\mathbf{p};s)$. The latter are very numerous and there may be insufficient observations available to estimate them econometrically. But Diewert's result tells us that we don't need to.

The quadratic mean of order *s* also includes the translog as a special case when *s* = 0; the Törnqvist is the corresponding superlative index. This can be seen by taking the limit as $s \to 0$ and applying de l'Hôpital's Rule. In the case where $s = 2$ the corresponding superlative index is the Fisher (Diewert, 1976). The Fisher and the Törnqvist are the forms most commonly used in empirical economics. The Fisher index is widely used by national statistical agencies, including those of the U.S.

[9] This follows from Shephard's Lemma which implies that the budget shares are given by $\partial \ln x / \partial \ln p_i$. These shares are independent of the level of utility and hence of expenditure when the expenditure function has the form of equation (4). So a doubling of expenditure with all prices held constant doubles the quantity purchased of every product.

As stated above, the quadratic mean of order s is only guaranteed to be a good approximation locally. As we move farther away from the point on which the approximation is based, it may cease to be a good one. The solution now is chaining. This means that we continue to believe that a quadratic mean of order s, with s assumed known, describes the data well, but the actual parameters can change over time. Eg, at time t the particular form given by (3) may apply, but at some other time r a related but different form may be a better approximation to consumer behaviour:

$$A'(\mathbf{p};s) = \left[\sum_{i=1}^{i=N}\sum_{j=1}^{j=N} b'_{ij} p_i^{s/2} p_j^{s/2}\right]^{1/s}, \; b'_{ij} = b'_{ji}, \forall i \neq j, s > 0 \tag{6}$$

where each b'_{ij} may differ from the corresponding b_{ij}. So in measuring the change in the Konüs price index between time t and $t+1$ equation (3) may apply, while from time r to time $r+1$ equation (6) may be better. Underlying preferences may be unchanged (the true utility function is unchanged), it's just that at some periods equation (3) may be a good approximation while at other periods equation (6) may be better. We don't need to know whether this is the case or not, because both sets of parameters are captured by the superlative index of equation (5). Hence chaining increases the flexibility of flexible functional forms by allowing parameters to change over time and this is consistent with preferences remaining unchanged.[10]

Hill (2006) has recently cast doubt on the optimistic conclusion that superlative indices solve the index number problem in the homothetic case. He argues that we have no good reason for picking one value of s over another and the value of the price index may be sensitive to the choice of s. He proves that as s is increased the value of the index approaches the geometric mean of the smallest and largest price relatives. Hence the index can be sensitive to outliers. He demonstrates this point using actual time series data for the US and cross country data for 43 countries and finds wide variations depending on the value of s. The spread between the largest and smallest values of a given index (for different values of s) often lies outside the Paasche-Laspeyres spread. However, there is not much variation in the indices as s increases from 0 (translog) to 2 (Fisher).

The optimistic conclusion can however be defended:

1. All Hill's comparisons are bilateral. He does not employ chain indices. But as argued above, chaining should substantially reduce the empirical uncertainty: the smaller the

[10] Diewert (1976) was well aware of this point: see his footnote 16.

change between adjacent years (or between countries), the closer will be the values of all superlative indices, ie they become increasingly insensitive to the choice of s.

2. If we adopt the economic approach (to which Hill is not necessarily committed), then the use of superlative indices requires that demand is homothetic. However unrealistic this is as a description of demand, it is the maintained hypothesis. But then theory implies that the true index must lie between the Paasche and the Laspeyres (Konüs, 1939, Deaton and Muellbauer, 1980b, chapter 7). So to be consistent with the maintained hypothesis, we should reject any value for the order s which produces a result outside the Paasche-Laspeyres spread. This again reduces the empirical uncertainty about the value of s.[11]

Unfortunately, the assumption of homotheticity is a very dubious one for demand. As argued earlier, there is overwhelming evidence from household surveys that income elasticities are not all equal to one. Economists have been somewhat readier to accept the assumption of constant returns to scale in the case of producers, but even so this assumption should ideally be tested. The next section therefore turns to the non-homothetic case.

3. Estimating a true cost-of-living index over time: the non-homothetic case

3.1 The Taylor series approach

In this section I consider the problem of how to estimate a true cost-of-living index over time when there are insufficient time series observations available to estimate the consumer's expenditure function.[12] This might be called the "large N, small T" problem: there are a large number of products but only a small number of time periods. This is the typical situation faced by national statistics agencies when for example estimating the consumer price index. Throughout this section I assume a single, representative consumer. In the next section this assumption will be relaxed.

[11] For Hill's the time series data, the maximum (absolute) Paasche-Laspeyres spread was 5% and the average one was 1.2%. For the cross-section data, the spread was much larger: 173.5% and 33.7% respectively. (I subtract 1 from his figures since he gives the ratio of Paasche to Laspeyres).

[12] The argument of this section is a generalisation of the one set out in Oulton (2008).

Let the share of product i in total expenditure at time t, if utility were fixed at the level of the base period b, be $s_i(t,b)$; ie the share is a function of the prices prevailing at t and the utility level at b. Applying Shephard's Lemma to the expenditure function, equation (1),

$$s_i(t,b) = \frac{\partial \ln E(t,b)}{\partial \ln p_i(t)}, \quad i = 1,\ldots,N \tag{7}$$

These can be called the hypothetical or *compensated* shares, the shares that would be observed if utility were held constant at some reference level (here, the level prevailing in period b), while prices followed their observed path. The actual, observed shares in period t are

$$s_i(t,t) = \frac{\partial \ln E(t,t)}{\partial \ln p_i(t)}, \quad i = 1,\ldots,N$$

Note that the compensated shares in the base period b, $s_i(b,b)$, are the same as the actual shares in that period.

By totally differentiating the Konüs price index of equation (2) with respect to time, we obtain

$$\frac{d \ln P^K(t,b)}{dt} = \sum_{i=1}^{i=N} \frac{\partial \ln E(t,b)}{\partial \ln p_i(t)} \frac{d \ln p_i(t)}{dt} = \sum_{i=1}^{i=N} s_i(t,b) \frac{d \ln p_i(t)}{dt} \tag{8}$$

So the level of the Konüs price index in some period T, relative to its level in the base period b, is:

$$\ln P^K(T,b) = \int_b^T \left[\sum_{i=1}^{i=N} s_i(t,b) \left(\frac{d \ln p_i(t)}{dt} \right) \right] dt \tag{9}$$

We see that the Konüs price index resembles a Divisia index but with the difference that the Konüs employs the compensated, not the actual, shares as weights (Balk, 2005; Oulton, 2008).[13]

In order to calculate the Konüs price index in practice, we seek a way of at least approximating the compensated shares, which cannot of course be directly observed (except for the $s_i(b,b)$ which are both the actual and the compensated share in period b). We can do this by expressing the actual shares $s_i(t,t)$ in terms of a Taylor series expansion of the

[13] Since it is a line integral, the Divisia index is in general path-dependent unless demand is homothetic, as its inventor Divisia (1925-26) was well aware; see Hulten (1973) for detailed discussion and Apostol (1957), chapter 10, for the underlying mathematics. But the Konüs price index is not path-dependent since by definition utility is being held constant along the path (Oulton, 2008).

compensated shares $s_i(t,b)$ around the point $\ln x = \ln E(t,b)$, ie holding prices constant at their levels at time *t* and varying real expenditure (utility):

$$s_i(t,t) = s_i(t,b) + \left(\frac{\partial s_i(\cdot,\cdot)}{\partial \ln E(\cdot,\cdot)} \right)_{\substack{\mathbf{p}=\mathbf{p}(t), \\ x=E(t,b)}} \cdot [\ln E(t,t) - \ln E(t,b)]$$

$$+ \frac{1}{2!} \left(\frac{\partial^2 s_i(\cdot,\cdot)}{\partial \ln E(\cdot,\cdot)^2} \right)_{\substack{\mathbf{p}=\mathbf{p}(t), \\ x=E(t,b)}} \cdot [\ln E(t,t) - \ln E(t,b)]^2 \tag{10}$$

$$+ \frac{1}{3!} \left(\frac{\partial^3 s_i(\cdot,\cdot)}{\partial \ln E(\cdot,\cdot)^3} \right)_{\substack{\mathbf{p}=\mathbf{p}(t), \\ x=E(t,b)}} \cdot [\ln E(t,t) - \ln E(t,b)]^3 + \ldots$$

Note that $\ln E(t,t) - E(t,b) = \ln[E(t,t)/E(t,b]$ is the log of the ratio of the expenditure needed to achieve the utility level of period *t* to the expenditure needed to achieve the level of period *b*, both evaluated at the prices of period *t*. In fact

$$\frac{E(t,t)}{E(t,b)} = \left[\frac{E(t,t)}{E(b,b)} \right] \cdot \left[\frac{E(b,b)}{E(t,b)} \right] = \frac{x(t,t)/x(b,b)}{P^K(t,b)} \tag{11}$$

where $x(v,v)$ is actual money expenditure at time *v* and we have used the definition of the Konüs price index in equation (2).

Now substitute (11) into (10) and solve for the compensated shares $s_i(t,b)$:

$$s_i(t,b) = s_i(t,t) - \eta_{i1}(t,b) \ln \left[\frac{x(t,t)/x(b,b)}{P^K(t,b)} \right] - \frac{\eta_{i2}(t,b)}{2!} \left\{ \ln \left[\frac{x(t,t)/x(b,b)}{P^K(t,b)} \right] \right\}^2$$

$$- \frac{\eta_{i3}(t,b)}{3!} \left\{ \ln \left[\frac{x(t,t)/x(b,b)}{P^K(t,b)} \right] \right\}^3 - \ldots, \tag{12}$$

$$i = 1, \ldots, N; \;\; t \in [0,T]$$

where to simplify the notation we have put

$$\eta_{ik}(t,b) = \left(\frac{\partial^k s_i(\cdot,\cdot)}{\partial \ln E(\cdot,\cdot)^k} \right)_{\substack{\mathbf{p}=\mathbf{p}(t), \\ x=E(t,b)}}, \quad k = 1, 2, \ldots; \; i = 1, \ldots, N \tag{13}$$

Equation (12) might not appear to take us very much further. But in fact it is the basis for a practical method of estimating the Konüs price index. The partial derivative $\eta_{i1}(t,b)$ is the semi-elasticity of the budget share of the *i*th product with respect to expenditure (real income), with prices held constant; it is evaluated at base year utility and at the prices of time *t*.

Suppose that this (and the higher order derivatives $\eta_{i2}(t,b)$, $\eta_{i3}(t,b)$, etc, that are required for a good approximation) were somehow known or could be estimated (see the next section on ways to do this). Then we could estimate the Konüs price index using equation (8) and (12). This is because these equations constitute a set of equations for $P^K(t,b)$, in which the compensated shares and the Konüs price index are the only unknowns; the actual shares $s_i(t,t)$, the money expenditures $x(t,t)$ and $x(b,b)$, and (by assumption) the semi-elasticities are all known.

The procedure to solve these equations is straightforward in principle. First, we need to take discrete approximations. Equations (12) must be understood to hold in discrete not continuous time, ie for $t = 0,1,...,T$. We must also decide how many terms in the Taylor series are required. If the utility function is quadratic in expenditure, then only the first two terms of the Taylor series are needed: see the next section. Equation (8) must be replaced by a discrete approximation, eg a chained Törnqvist (P^T) or chained Fisher formula (P^F).

Let us define the following chained, *compensated* index numbers. Each index number is for time t relative to time $t-1$, with utility held constant at the level of period *b*.

Compensated Törnqvist:

$$\ln P^T(t,t-1,b) = \sum_{i=1}^{i=N} \left(\frac{s_i(t,b) + s_i(t-1,b)}{2} \right) \ln \left(\frac{p_i(t)}{p_i(t-1)} \right) \tag{14}$$

Compensated Laspeyres:

$$P^L(t,t-1,b) = \sum_{i=1}^{i=N} s_i(t-1,b) \frac{p_i(t)}{p_i(t-1)} \tag{15}$$

Compensated Paasche:[14]

$$P^P(t,t-1,b) = \left[\sum_{i=1}^{i=N} s_i(t,b) \frac{p_i(t-1)}{p_i(t)} \right]^{-1} \tag{16}$$

Compensated Fisher:

$$P^F(t,t-1,b) = [P^L(t,t-1,b) \cdot P^P(t,t-1,b)]^{1/2} \tag{17}$$

Each of these index numbers is defined in the same way as its empirical counterpart, except that compensated, not actual, shares are used. The natural choices for discrete approximations to the continuous Konüs price index are either the compensated Törnqvist, equation (14), or the compensated Fisher, equation (17).

[14] The formula for the Paasche is not the usual one but is mathematically equivalent to the usual one.

Since utility is being held constant at its level in period b, the true index is bounded by the compensated Laspeyres and the compensated Paasche:

$$P^L(t,t-1,b) \geq P^K(t,b)/P^K(t-1,b) \geq P^P(t,t-1,b) \tag{18}$$

This follows from the well-known Konüs (1939) inequalities (the proof is in the Annex). The Paasche-Laspeyres spread, calculated using the compensated shares, can be used as a check on the accuracy of whatever index number formula is adopted.[15]

Equations (12) now constitute a system of $(N-1)(T+1)$ independent equations since the N shares sum to one in each period.[16] Together with (8), this system can be solved iteratively:

1. Start with an initial guess at $P^K(t,b)$: this could be derived as a chained Törnqvist or chained Fisher index which uses actual not compensated shares.

2. Substitute this estimate of $P^K(t,b)$ into (12) to get estimates of the compensated shares for each of $N-1$ products and for each of $T+1$ time periods; the share of the Nth product can be derived as a residual.

3. Use these estimates of the compensated shares to obtain a new estimate of $P^K(t,b)$ from either of the two discrete approximations to (8), the Törnqvist (equation (14)) or the Fisher (equation (17)).[17]

4. Check whether the estimate of $P^K(t,b)$ has converged. If not, return to step 2.

So given knowledge of the η_{ik} up to the required order, we can estimate the Konüs price index. Estimating the η_{ik} may still seem a difficult task but notice that only the response of demand to changes in real income needs to be known, not the response to price changes. This is a very significant reduction in the complexity of the task empirically. To make further progress we turn now to consider systems of demand which are consistent with economic theory and also seem capable of fitting the data reasonably well.

---

15 Of all superlative index numbers, only the Fisher is guaranteed to lie within the Laspeyres-Paasche spread (Hill, 2006).

16 The actual shares of course sum to one and since they derive from the expenditure function so do the compensated shares: see equation (7).

17 In step 3 of the algorithm it is assumed that the observations are arranged in the natural time order. See below for a refinement.

3.2 Demand systems

The PIGLOG demand system, introduced by Muellbauer (1976) (see also Deaton and Muellbauer (1980a and 1980b, chapter 3)) has found wide application empirically; an example of the PIGLOG is the AIDS system. The PIGLOG expenditure function is:

$$\ln x = \ln A(\mathbf{p}) + B(\mathbf{p}) \ln u \tag{19}$$

Here $A(\mathbf{p}) \geq 0$ and $B(\mathbf{p}) > 0$ (non-satiation). Also, $A(\mathbf{p})$ is assumed homogeneous of degree 1 and $B(\mathbf{p})$ homogeneous of degree 0 in prices. From Shephard's Lemma, the share equations for this system are:

$$s_i = \frac{\partial \ln A(\mathbf{p})}{\partial \ln p_i} + \frac{\partial B(\mathbf{p})}{\partial \ln p_i} \ln u$$

or, using (19)

$$s_i = \frac{\partial \ln A(\mathbf{p})}{\partial \ln p_i} + \frac{\partial \ln B(\mathbf{p})}{\partial \ln p_i} \ln\left[x / A(\mathbf{p})\right] \tag{20}$$

(The homothetic case is where $\partial \ln B(\mathbf{p}) / \partial \ln p_i = 0, \forall i$). The AIDS system specifies that $B(\mathbf{p}) = \prod_k p_k^{\beta_k}$, $\sum_k \beta_k = 0$, in which case the coefficient on $\ln[x / A(\mathbf{p})]$ in the share equations is β_i, a constant independent of prices.

However, a linear relationship between the share and the log of deflated expenditure as in (20) does not fare well empirically (Banks *et al.*, 1997; Blow *et al.*, 2004; Oulton, 2008) and it is found necessary to add a squared term in the log of deflated expenditure. Lewbel (1991) defined the rank of a demand system to be the dimensions of the space spanned by its Engel curves. Exactly aggregable demand systems are those which are linear in functions of *x*. Gorman (1981) proved that the maximum possible rank of any exactly aggregable demand system is 3. The empirical evidence on Engel curves indicates that observed demands are at least rank 3. Theorem 1 of Banks *et al.* (1997) states that all exactly aggregable, rank 3, demand systems which just add a differentiable function of deflated expenditure to equation (20) are derived from a utility function of the form

$$\begin{aligned} \ln u &= \left\{ \left[\frac{\ln x - \ln A(\mathbf{p})}{B(\mathbf{p})} \right]^{-1} + \lambda(\mathbf{p}) \right\}^{-1} \\ &= \frac{\ln[x / A(\mathbf{p})]}{B(\mathbf{p}) + \ln[x / A(\mathbf{p})] \lambda(\mathbf{p})} \end{aligned} \tag{21}$$

where $\lambda(\mathbf{p})$ is a differentiable, homogeneous function of degree zero in prices **p**. The corresponding expenditure function is:

$$\ln x = \ln A(\mathbf{p}) + \frac{B(\mathbf{p})}{1-\lambda(\mathbf{p})\ln u}\ln u \tag{22}$$

(This reduces to the PIGLOG system (19) when $\lambda(\mathbf{p}) = 0$).

Applying Shephard's Lemma, the budget shares in this demand system are:

$$s_i = \frac{\partial \ln A(\mathbf{p})}{\partial \ln p_i} + \frac{\ln u}{1-\lambda(\mathbf{p})\ln u}\frac{\partial B(\mathbf{p})}{\partial \ln p_i} + \frac{1}{B(\mathbf{p})}\left[\frac{B(\mathbf{p})\ln u}{1-\lambda(\mathbf{p})\ln u}\right]^2 \frac{\partial \lambda(\mathbf{p})}{\partial \ln p_i}$$

Hence from (22)

$$s_i = \frac{\partial \ln A(\mathbf{p})}{\partial \ln p_i} + \frac{\ln[x/A(\mathbf{p}]}{B(\mathbf{p})}\frac{\partial B(\mathbf{p})}{\partial \ln p_i} + \frac{[\ln[x/A(\mathbf{p})]]^2}{B(\mathbf{p})}\frac{\partial \lambda(\mathbf{p})}{\partial \ln p_i} \tag{23}$$

In equation (12) above we found a Taylor series expansion for the compensated shares which involved the semi-elasticity of the shares with respect to real income, $\partial s_i / \partial \ln E$, and higher order derivatives, $\partial^2 s_i / \partial \ln E^2$, etc. Now from (23) we get that

$$\begin{aligned} \frac{\partial s_i}{\partial \ln x} &= \frac{\partial \ln B(\mathbf{p})}{\partial \ln p_i} + \frac{2}{B(\mathbf{p})}\frac{\partial \lambda(\mathbf{p})}{\partial \ln p_i}\ln[x/A(\mathbf{p})] \\ \frac{\partial^2 s_i}{\partial [\ln x]^2} &= \frac{2}{B(\mathbf{p})}\frac{\partial \lambda(\mathbf{p})}{\partial \ln p_i} \end{aligned} \tag{24}$$

and higher order derivatives are zero.

These derivatives have to be evaluated when $x = E(t,b)$. The simplest way to do this is to adopt the normalisation that $\ln[x(b,b)/A(\mathbf{p}(b)] = 0$. From (22)

$$\ln x(b,b) = \ln A(\mathbf{p}(b)) + \frac{B(\mathbf{p}(b))\ln u(b)}{1-\lambda(\mathbf{p}(b))\ln u(b)} \tag{25}$$

Now choose monetary and quantity units so that $x(b,b) = A(\mathbf{p}(b)$. This is always possible since $A(\mathbf{p})$ depends only on prices while $x = \sum_i p_i q_i$ depends on both prices and quantities. For example, suppose that x is initially double $A(\mathbf{p})$ at time b. Then increase all quantity units by 100% and increase all prices correspondingly by 100%.[18] This doubles $A(\mathbf{p})$ while leaving x unchanged. Then under this normalisation (25) implies that

[18] Equality between $A(\mathbf{p}(b))$ and $x(b,b)$ can be achieved by an appropriate change in the monetary unit. Suppose that, after normalising prices to one in the reference year, $A(\mathbf{p}(b))$ is

$$\ln u(b) = 0$$

It then follows also from (22) that

$$\ln x(t,b) = \ln A_b(\mathbf{p}(t)) + \frac{B(\mathbf{p}(t))\ln u(b)}{1-\lambda(\mathbf{p}(t))\ln u(b)} = \ln A_b(\mathbf{p}(t)) \tag{26}$$

From now on, we write $A_b(\mathbf{p})$ rather than just $A(\mathbf{p})$, to mark the fact that prices are now scaled by factors specific to period *b*. In general, this normalisation changes matrix *A* (see below for the change in the QAIDS case).

This last finding suggests that we can interpret $A_b(\mathbf{p})$ as the Konüs price index with base period *b*. More formally, using the definition of the Konüs price index, equation (2), and equation (26), for the generalised PIGLOG we find that:

$$\ln P^K(t,b) = \ln E(t,b) - \ln E(b,b) = \ln A_b(\mathbf{p}(t) - \ln A_b(\mathbf{p}(b)) \tag{27}$$

ie $P^K(t,b) = A_b(\mathbf{p}(t)) / A_b(\mathbf{p}(b))$. In other words, with this normalisation the Konüs price index is measured by the homothetic part of the expenditure function $A_b(\mathbf{p})$, so $\left[x(t,t)/x(b,b)\right] / \left[A_b(\mathbf{p}(t))/A_b(\mathbf{p}(b))\right]$ measures real income relative to its level in period *b*.

We can now use these results to evaluate the derivatives in (24) at the point $x = E(t,b),\ \mathbf{p} = \mathbf{p}(t)$:

$$\eta_{i1}(t,b) = \left[\frac{\partial s_i}{\partial \ln x}\right]_{\substack{\mathbf{p}=\mathbf{p}(t)\\ x=E(t,b)}} = \frac{\partial \ln B(\mathbf{p}(t))}{\partial \ln p_i(t)} + \frac{2}{B(\mathbf{p}(t))}\frac{\partial \lambda(\mathbf{p}(t))}{\partial \ln p_i(t)} \ln\left[\frac{x(t,b)}{A_b(\mathbf{p}(t))}\right]$$

$$= \frac{\partial \ln B(\mathbf{p}(t))}{\partial \ln p_i(t)}$$

using (26) and

$$\eta_{i2}(t,b) = \left[\frac{\partial^2 s_i}{\partial[\ln x]^2}\right]_{\substack{\mathbf{p}=\mathbf{p}(t)\\ x=E(t,b)}} = \frac{2}{B(\mathbf{p}(t))}\frac{\partial \lambda(\mathbf{p}(t))}{\partial \ln p_i(t)}$$

Substituting these results into (12) we obtain

---

now a multiple $H_b > 0$ of $x(b,b)$. Then define a new monetary unit as H_b times the old one. This leaves the normalised prices, and so also $A(\mathbf{p}(b))$, unchanged but multiplies the value of expenditure by H_b. See the Annex for more detail.

$$s_i(t,b) = s_i(t,t) - \frac{\partial \ln B(\mathbf{p}(t))}{\partial \ln p_i(t)} \ln\left[\frac{x(t,t)/x(b,b)}{P^K(t,b)}\right]$$

$$-\frac{1}{B(\mathbf{p}(t))}\frac{\partial \lambda(\mathbf{p}(t))}{\partial \ln p_i(t)}\left\{\ln\left[\frac{x(t,t)/x(b,b)}{P^K(t,b)}\right]\right\}^2$$

and this Taylor series expansion is not an approximation but exact for the generalised PIGLOG.

To illustrate how much this simplifies the task of estimating the Konüs price index, I turn now to a specific example, the Quadratic Almost Ideal Demand System (QAIDS).[19]

3.3 The AIDS and QAIDS cases

The QAIDS is an example of a generalised PIGLOG system. Banks *et al.* (1997) specify that

$$B(\mathbf{p}) = \prod_k p_k^{\beta_k}, \ \sum_k \beta_k = 0 \tag{28}$$

as in the simpler AIDS and

$$\lambda(\mathbf{p}) = \sum_k \lambda_k \ln p_k, \ \sum_k \lambda_k = 0 \tag{29}$$

Under this specification,

$$\frac{\partial \ln B(\mathbf{p})}{\partial \ln p_i} = \beta_i$$

$$\frac{\partial \lambda(\mathbf{p})}{\partial \ln p_i} = \lambda_i$$

The share equations are then[20]

---

[19] Lewbel and Pendakur (2009) have recently proposed a new demand system, the Exact Affine Stone Index (EASI) system. This has all the advantages of the generalised PIGLOG (and of the QAIDS) while allowing Engel curves to be still more flexible, eg polynomials of cubic or higher order. In principle the method developed here could be applied to the EASI system as well. However, I have not been able to develop tractable expressions for the derivatives of the share equations with respect to log expenditure (the η_{ik}). From the point of view of the present paper, the EASI system suffers from the disadvantage that exact aggregation does not hold. This does not matter when the system is fitted to individual data but does when fitted to aggregate data: see section 4 for discussion of aggregation over consumers who may differ in income and in other ways.

[20] The coefficient on $\ln[x/A(\mathbf{p})]$ in (30) is independent of prices while that on $\ln[x/A(\mathbf{p})]^2$ is not. Banks *et al.* (1997) show in their Corollary 2 that the two coefficients cannot both be independent of prices if the system is rank 3.

$$s_i = \frac{\partial \ln A(\mathbf{p})}{\partial \ln p_i} + \beta_i \ln\left[\frac{x}{A(\mathbf{p})}\right] + \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left\{\ln\left[\frac{x}{A(\mathbf{p})}\right]\right\}^2 \tag{30}$$

At this point we do not need to specify a functional form for $A(\mathbf{p})$ but it is worth noting that the AIDS form

$$\ln A(\mathbf{p}) = \alpha_0 + \sum_i \alpha_i \ln p_i + (1/2)\sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j,\ \sum_i \alpha_i = 1, \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0 \tag{31}$$

would lead to share equations of the form:

$$s_i = \alpha_i + \sum_{j=1}^{j=N} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{x}{A(\mathbf{p})}\right] + \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left\{\ln\left[\frac{x}{A(\mathbf{p})}\right]\right\}^2 \tag{32}$$

In equation (12) above we found a Taylor series expansion for the compensated shares which involved the semi-elasticity of the shares with respect to real income, $\partial s_i / \partial \ln E$, and higher order derivatives, $\partial^2 s_i / \partial \ln E^2$, etc. Now using (30), we get that

$$\begin{aligned} \frac{\partial s_i}{\partial \ln x} &= \beta_i + \frac{2\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \ln\left[\frac{x}{A(\mathbf{p})}\right] \\ \frac{\partial^2 s_i}{\partial [\ln x]^2} &= \frac{2\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \end{aligned} \tag{33}$$

and higher order derivatives are zero.

We can now use these results to evaluate the derivatives in (33) at the point $x = E(t,b),\ \mathbf{p} = \mathbf{p}(t)$ after applying the normalisation of equation (26):

$$\eta_{i1}(t,b) = \left[\frac{\partial s_i}{\partial \ln x}\right]_{\substack{\mathbf{p}=\mathbf{p}(t) \\ x=E(t,b)}} = \beta_i$$

$$\eta_{i2}(t,b) = \left[\frac{\partial^2 s_i}{\partial [\ln x]^2}\right]_{\substack{\mathbf{p}=\mathbf{p}(t) \\ x=E(t,b)}} = \frac{2\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}(t)}$$

Substituting these results into (12) we obtain

$$s_i(t,b) = s_i(t,t) - \beta_i \ln\left[\frac{x(t,t)/x(b,b)}{P^K(t,b)}\right] - \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left\{\ln\left[\frac{x(t,t)/x(b,b)}{P^K(t,b)}\right]\right\}^2, \tag{34}$$

$$i = 1,2,\ldots,N; \quad t = 0,1,\ldots,T-1$$

and this Taylor series expansion is not an approximation but exact for the generalised PIGLOG with the specification of (28) and (29).

Therefore in order to implement the procedure outlined above for estimating the Konüs price index, we need to estimate only the N β_i parameters and the N λ_i parameters; in both cases only $N-1$ of these are independent because these coefficients each sum to zero across the products. That is, $2(N-1)$ parameters in total need to be estimated or just two per share equation. These parameters determine the consumer's response to changes in real income. We do *not* need to estimate the much more numerous parameters which determine the response to price changes. This is a huge reduction in the difficulty of the task.

There are now two ways to proceed. Either we can use estimates of the β_i and λ_i parameters that have been derived independently or we can seek to estimate them from the price and quantity data that are employed to calculate conventional index numbers. The response to income changes can be estimated from cross section data since prices can often be assumed to be the same for all households (see eg Blow *et al.* (2004)). But cross section estimates may not be available or, if they are, the product classification may be different. So there is interest in seeing whether these parameters can be estimated from using just the aggregate price and quantity data. I show how this can be done in the next subsection.

3.4 The estimation procedure

Even if we need only the income response parameters, how can we estimate these while avoiding estimating all the other parameters of the system at the same time? After all, if we just estimate the share equations with the price variables omitted then our estimates of the income response will undoubtedly be biased, since relative prices and real incomes are likely to be correlated over time and (and across countries). The answer is to collapse the $N-1$ relative prices in the system into a smaller number of variables using principal components.[21] We can collapse the relative price data into (say) M principal components, where $M < N-1$ is to be chosen empirically.

To implement this idea, start by applying the normalisation of equation (26). This implies that $x(b,b) = A_b(\mathbf{p}(b))$, so, making use of (27), the share equations (30) can be written as:

[21] See Johnson and Wichern (2002) for a textbook exposition of principal components.

$$s_i(t,t) = \frac{\partial \ln A_b(\mathbf{p}(t))}{\partial \ln p_i(t)} + \beta_i \ln\left[\frac{x(t,t)/x(b,b)}{A_b(\mathbf{p}(t))/A_b(\mathbf{p}(b))}\right] + \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left\{\ln\left[\frac{x(t,t)/x(b,b)}{A_b(\mathbf{p}(t))/A_b(\mathbf{p}(b))}\right]\right\}^2$$

$$= \frac{\partial \ln A_b(\mathbf{p}(t))}{\partial \ln p_i(t)} + \beta_i \ln\left[\frac{x(t,t)/x(b,b)}{P^K(t,b)}\right] + \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left\{\ln\left[\frac{x(t,t)/x(b,b)}{P^K(t,b)}\right]\right\}^2$$

These equations can be written in a form suitable for econometric estimation as:

$$s_i(t,t) = \alpha_i^b + \sum_{k=1}^{M} \theta_{ik} PC_k(t) + \beta_i z(t,b) + \lambda_i y(t,b) + \varepsilon_i(t), \qquad i = 1,\dots,N;\ t = 0,\dots,T \tag{35}$$

where α_i^b is the base-year-dependent constant term ($\sum_i \alpha_i^b = 1$); $PC_k(t)$ is the kth principal component of the $N-1$ relative prices; the θ_{ik} are coefficients subject to the cross-equation restrictions $\sum_i \theta_{ik} = 0,\ \forall k$; ε_{it} is the error term; and we have put $z(t,b) = \ln[x(t,t)/P^K(t,b)]$ as before and also $y(t,b) = [z(t,b)]^2 / \prod_k p_k^{\beta_k}(t)$. The presence of the principal components in equation (35) means that the estimates of the coefficients on *z* and *y* need not be biased as they would be if prices were simply omitted.[22]

We have now reduced the problem to estimating a system of $N-1$ independent equations, each of which contains only $M+3$ coefficients — the θ_{ik} (*M* in number), α_i, β_i and λ_i.[23] The success of this strategy will depend on whether the variation in relative prices can be captured by a fairly small number of principal components — small that is in relation to the number of time series observations, $T+1$. This is obviously an empirical matter. At one extreme, if there is little or no correlation between the prices over time (or space), then the use of principal components yields no benefit. At the other extreme, suppose that the demand system is specified in terms of the logs of prices (as in the AIDS and QAIDS) and that all relative prices are just loglinear time trends, though the growth rate varies between prices. The evolution of relative prices can be written as:

---

22 The empirical flexibility of equation (35) could be increased by adding cubic and higher order terms in $z(t,b)$. (The coefficients on these additional terms must be constrained to sum to zero across products). The implied expenditure function could not now be written down in closed form but the share equations extended in this way could be regarded as polynomial approximations to the exact ones. However, in the presence of cubic and higher order terms the property of exact aggregation would no longer hold, making it hard to interpret the results in terms of individual welfare. See the next section for more on aggregation.

23 This is not quite true since all the β_i appear in each equation via the denominator of *y*. We can handle this by an iterative procedure: see below.

$$\ln[p_j(t)/p_1(t)] = \mu_j t, \quad j = 2,...,N$$

where the μ_j are the growth rates and the first product is taken as the numeraire. Then in the share equations (32) the price effects are

$$\sum_{j=1}^{N} \gamma_{ij} \ln p_j(t) = \sum_{j=2}^{N} \gamma_{ij} \ln[p_j(t)/p_1(t)] = t\left[\sum_{j=2}^{N} \gamma_{ij}\mu_j\right] = \delta_i t \text{, say.}$$

(Here we have used the fact that $\sum_j \gamma_{ij} = 0$). In this case the effect of relative prices is captured entirely by a time trend, with a different coefficient in each share equation (subject to the cross-equation restriction that $\sum_i \delta_i = 0$). So just one principal component captures the whole variation in relative prices (ie in this case $M = 1$). This is an extreme case and in practice we must expect that more than one principal component will be required to capture the variation in relative prices.[24]

The specification of the principal components depends on the demand system chosen. If we chose the AIDS (and QAIDS) form for $A(\mathbf{p})$, then it would be natural to estimate the principal components in terms of log relative prices, eg $\ln(p_j / p_1),\ j = 2,...,N$, taking the first product as the numeraire. Alternatively, we might use the normalised quadratic of Diewert and Wales (1988), in which case the principal components would be estimated in terms of relative prices (not in logs).

In estimating equation (35) econometrically, it is straightforward to impose the adding-up and homogeneity restrictions on the coefficients; homogeneity is imposed by using relative prices and adding-up is imposed by cross-section restrictions on the coefficients (these restrictions are automatically imposed by OLS though the latter is not necessarily the best method). But there is one loss from using principal components: we can no longer impose the symmetry restrictions.[25]

Equation (35) is nonlinear in the parameters of interest, since to measure both z and y correctly it is necessary to know the Konüs price index, the object of the whole exercise; in

[24] In Oulton (2008) I applied the method to 70 products covering the whole of the U.K.'s Retail Prices Index over 1974-2004. I found that six principal components were sufficient to capture 97.8% of the variation in the 69 log relative prices.

[25] For example, suppose that $N = 3$ and that the special case of all relative prices changing at constant rates applies. Then, dropping the third equation, taking the first product as the numeraire, and imposing all the constraints, the relationship between the δ_i and the γ_{ij} is as follows: $\delta_1 = \gamma_{12}\mu_2 - (\gamma_{11} + \gamma_{12})\mu_3$, $\delta_2 = \gamma_{22}\mu_2 - (\gamma_{12} + \gamma_{22})\mu_3$. These relationships imply no further restrictions on δ_1 and δ_2. So we cannot test whether $\gamma_{12} = \gamma_{21}$.

addition, to measure y we also need to know all the β_i and λ_i. The solution is an iterative process, similar to the one described in the previous section. Here the unknown parameters, the β_i and λ_i, are estimated jointly with the compensated shares and the Konüs price index. The system consists of equations (34), (35) and the equation for the Konüs price index, either equation (14) if we use a compensated Törnqvist to approximate the Konüs or equation (17) if we use a compensated Fisher. The iterative process for some particular choice of the base period is as follows:

1. Obtain initial estimates of the Konüs price index $P^K(t,b)$ and of the β_i and λ_i coefficients. An initial estimate of $P^K(t,b)$ can be obtained from equation (14) or equation (17) by using actual instead of compensated shares (ie replace $s_i(t,b)$ by $s_i(t,t)$ in the formulas). And for an initial estimate of the β_i and λ_i, set $\beta_i = \lambda_i = 0, \forall i$.
2. Derive estimates of $z(t) = \ln[x(t) / P^K(t,b)]$ and of $y(t) = [z(t,b)]^2 / \prod_k p_k^{\beta_k}$, using the latest estimates of $P^K(t,b)$ and of the β_i. Using these new estimates of z and y, estimate equation (35) econometrically, to obtain new estimates of the β_i and λ_i.
3. Using the new estimates of the β_i and λ_i, estimate the compensated shares from equation (34). Then use the compensated shares to derive a new estimate of the Konüs price index $P^K(t,b)$ from equation (14) or equation (17).
4. If the estimate of the Konüs price index has changed by less than a preset convergence condition, stop. If not, go back to step 2.

Finally, the estimates of the β_i and λ_i produced by the algorithm above can be plugged into the simpler algorithm of section 3.1 to generate Konüs price indices for any other base year.

3.6 Comparisons across space

The analysis carries over unchanged to the problem of estimating a cost of living index and hence the standard of living across countries at a point in time.[26] The solution for the Konüs

[26] See Hill (1997) for a survey of methods of making international comparisons. Caves *et al.* (1982) have applied chained superlative index numbers to cross-country comparisons. Hill (2004) also estimates a chain superlative index but employs the minimum-spanning tree

price index given by equations (12) and (9) can be applied directly in the cross-country context. Initially we must imagine a continuum of countries indexed by t just as in section 3 we imagined a continuum of time periods. Then we consider discrete approximations, eg as before equation (9) can be approximated by either (14) or (17).

One problem which is often said to arise in the cross-country but not the intertemporal context is that, unlike time, countries have no natural order. In the present case this objection does not apply. Here the natural order for countries is the ranking by real income (or real expenditure) per capita. Adopting this order minimises the gap between country t and country $t-1$ and so should improve the discrete approximation. It is true that the rank order is not known for certain in advance, since the whole point of the exercise is to estimate the true standard of living. This suggests a refinement to the algorithm: at each step, re-order the countries (time periods) so as to put them in rank order of real expenditure per capita (where "real" means deflated by the algorithm's latest estimate of the Konüs price index). Alternatively, the ordering of countries could be determined by the minimum-spanning-tree method suggested by and implemented on cross-country data by of Hill (1999). Then the links in the chain would be selected so as to minimise the (compensated) Paasche-Laspeyres spread.[27]

4. Extensions to the basic analysis

The preceding section 3 offered a solution to the problem of estimating a true cost-of-living index over time for a single representative consumer. In this section, I consider two extensions to the analysis. First, I consider the effect of relaxing the assumption of a single representative consumer. I now assume that the aggregate data is generated by heterogeneous consumers who differ in income. If the degree of inequality were constant the preceding analysis could stand unchanged. This may or may not be a reasonable approximation in a time series context over a few decades. But in a cross-country context the assumption is certainly problematic: countries differ widely in the extent of inequality (Anand and Segal,

---

approach to find the best links in the chain. Neary (2004) employed the World Bank's 1980 PPPs for 60 countries and 11 commodity groups to estimate a QAIDS; he then derived a measure of real GDP per capita for the 60 countries. The World Bank's current methodology for deriving PPPs at the aggregate level is set out in World Bank (2008).

[27] Hill (2004) uses a different criterion, namely minimising a dissimilarity index suggested by Diewert (2002), but this seems less appropriate in the present context.

2008). So we need to extend our framework to encompass this. Second, I consider aggregation over different types of household.

4.1 Aggregation over rich and poor consumers

Let the population be composed of *G* groups. The groups are assumed to be of equal size (eg percentiles, deciles or quintiles), with the first group being the poorest and the *G*th group the richest. The fraction of households in each group is then $1/G$. Let x_g be mean expenditure per household in the *g*th group. Within a group, each household's expenditure is the same, namely the group mean. The share of product *i* in the expenditure of the *g*th group, s_{ig} ,is then

$$s_{ig} = \frac{p_i q_{ig}}{x_g}$$

where q_{ig} is the quantity per capita of the *i*th product purchased by each member of the *g*th group. The share of the *i*th product in aggregate expenditure is therefore

$$s_i = \frac{p_i q_i}{x} = \frac{\sum_{g=1}^{g=G} p_i q_{ig}}{Gx} = \sum_{g=1}^{g=G} \left[\frac{x_g}{Gx} \frac{p_i q_{ig}}{x_g} \right] = \sum_{g=1}^{g=G} w_g s_{ig} \tag{36}$$

where w_g is the share of the *g*th group in aggregate expenditure:

$$w_g = \frac{x_g}{Gx}, \quad \sum_{g=1}^{g=G} w_g = 1 \tag{37}$$

We assume that preferences have the Ernest Hemingway property: the rich are different from the poor but only because the rich have more money.[28] So the parameters of the expenditure function are the same for all households. All consumers are assumed to face the same prices. So from (30) and adopting the QAIDS formulation, the share of the *i*th product in expenditure by the *g*th group is:

$$s_{ig} = \alpha_i + \sum_{j=1}^{j=N} \gamma_{ij} \ln p_j + \beta_i \ln \left[\frac{x_g}{A(\mathbf{p})} \right] + \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \left\{ \ln \left[\frac{x_g}{A(\mathbf{p})} \right] \right\}^2$$

where $A(\mathbf{p})$ takes the AIDS form of equation (31). Using (36), the aggregate share equations are weighted averages of the underlying equations for each group:

---

[28] The well-known dialogue runs as follows. Fitzgerald: "The rich are different from us, Ernest". Hemingway: "Yes, Scott, they have more money than we do".

$$s_i = \sum_{g=1}^{g=G} w_g s_{ig} = \alpha_i + \sum_{j=1}^{j=N} \gamma_{ij} \ln p_j + \beta_i \sum_{g=1}^{g=G} w_g \ln x_g - \beta_i \ln A(\mathbf{p})$$

$$+ \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \left[\sum_{g=1}^{g=G} w_g (\ln x_g)^2 - 2 \ln A(\mathbf{p}) \sum_{g=1}^{g=G} w_g \ln x_g + [\ln A(\mathbf{p})]^2 \right] \tag{38}$$

The difference between this and our previous equation (30) is that instead of the log of aggregate expenditure per capita, $\ln x = \ln\left[\sum_{g=1}^{g=G} x_g / G\right]$, appearing on the right hand side, we now have the share-weighted average of log expenditure per capita in each group, $\sum_{g=1}^{g=G} w_g \ln x_g$; and instead of $(\ln x)^2$, we now have $\sum_{g=1}^{g=G} w_g (\ln x_g)^2$. The relationship between $\sum_{g=1}^{g=G} w_g \ln x_g$ and $\ln x$ is, from (37),

$$\sum_{g=1}^{g=G} w_g \ln x_g = \sum_{g=1}^{g=G} w_g \ln(w_g G x) = \sum_{g=1}^{g=G} w_g \ln w_g + \ln G + \ln x$$

The first term on the right hand side, $\sum_g w_g \ln w_g$, is the negative of entropy (ignoring an unimportant scale constant); it was suggested as a measure of inequality by Theil (1967), chapter 4. Substituting this into (38), we find after some manipulation (see the Annex) that

$$s_i = [\alpha_i + \beta_i \ln G] + \sum_{j=1}^{j=N} \gamma_{ij} \ln p_j$$

$$- \beta_i I + \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \left[J - 2I \ln G + (\ln G)^2 \right] \tag{39}$$

$$+ \left(\beta_i + \frac{2\lambda_i (\ln G - I)}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \right) \ln\left[\frac{x}{A(\mathbf{p})}\right] + \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \left\{ \ln\left[\frac{x}{A(\mathbf{p})}\right] \right\}^2$$

where we have set $I = -\sum_{g=1}^{g-G} w_g \ln w_g$ and $J = \sum_{g=1}^{g-G} w_g (\ln w_g)^2$. In the case of a perfectly equal distribution (when $w_g = 1/G$), note that $I = \ln G$ and $J = (\ln G)^2$, so that (39) then reduces back down to the original QAIDS formulation, equation (32). The constant term in (39) is now $\alpha_i + \beta_i \ln G$ which continues to sum to one across products. Compared to (32), there are two additional variables in (39), entropy (*I*) and a related statistic (*J*), though no additional parameters. These additional variables may help to explain changes in shares, to the extent that inequality varies either over time or across countries. Note too that in the simpler AIDS case (ie when all the λ_i are zero), equation (39) simplifies to

$$s_i = [\alpha_i + \beta_i \ln G] + \sum_{j=1}^{j=N} \gamma_{ij} \ln p_j - \beta_i I + \beta_i \ln\left[\frac{x}{A(\mathbf{p})}\right] \tag{40}$$

which contains just one additional variable (I).[29]

The upshot is that the QAIDS can be parsimoniously extended to capture the effect of income inequality. The additional empirical requirement is fairly modest: we need to know the shares of different groups in aggregate expenditure, at a reasonable level of detail.

4.2 Aggregation over different household types

Suppose there are a set of H characteristics that influence demand, in addition to income and prices. These could include household characteristics such as number of children, average age, and educational level, and also environmental characteristics such as climate. Now the share equations of the QAIDS for the gth income group could be written as:

$$s_{ig} = \alpha_i + \sum_{j=1}^{j=N} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{x_g}{A(\mathbf{p})}\right] + \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \left\{\ln\left[\frac{x_g}{A(\mathbf{p})}\right]\right\}^2 + \sum_{n=1}^{n=H} \theta_{ih} K_{hg} \tag{41}$$

where K_{hg} is the level of the hth characteristic in the gth group; I assume that each household in the gth group has the same level of each of the K_{hg} as all the other households in that group (this entails no loss of generality if there is only one household in each group). The θ_{in} coefficients must satisfy the adding-up restrictions:

$$\sum_{i=1}^{i=H} \theta_{in} = 0, \quad n = 1, 2, ..., H$$

(At some cost to parsimony, the model could be extended by interacting the characteristic variables with income). Again, underlying preferences are assumed to be the same but people's situations differ for various reasons, in the spirit of Stigler and Becker (1977): at the same incomes and prices, people in cold climates buy more winter clothes. We can aggregate equation (41) over the income groups to obtain the same result as (39), but with an additional term:

$$+\sum_{h=1}^{h=H} \theta_{ih} K_h$$

where $K_h = \sum_{g=1}^{g=G} w_g K_{hg}$. Now K_h is a weighted average of the level of the hth characteristic in a particular country (time period). The only difficulty from an empirical point of view is

[29] The role of Theil's inequality measure, entropy (I), was discussed in Deaton and Muellbauer (1980b) chapter 6, section 6.2. They derived a result equivalent to (40).

that it is an income-weighted, not a population-weighted, average. So for example if the rich have fewer children than the poor, then using the mean number of children per household as a measure would be a misspecification when estimating share equations from aggregate data.

5. Cost functions: estimating input-biased scale economies and technical progress

In this section I look at the parallel problem of estimating an input price index and technical progress when the cost function is not homothetic. Now both economies of scale and technical progress may be input-biased. I assume that the typical firm is a price taker in input markets and wishes to minimise costs. We can write the cost function in general as:

$$x = C(\mathbf{p}, Y, t) \tag{42}$$

Here output (*Y*) plays the role of utility in the expenditure function. While formally this makes no difference, there is a big difference empirically since output is objectively and directly measurable (at least in principle) while utility is only indirectly measurable. The presence of time (*t*) as an indicator of technical progress in the cost function also has no counterpart in the theory of demand.[30]

By analogy with equation (22), we can use a generalised PIGLOG formulation:

$$\ln x = \ln C(\mathbf{p}, Y, t) = \ln A(\mathbf{p}) + \frac{B(\mathbf{p}) \ln Y}{1 - \lambda(\mathbf{p}) \ln Y} + \beta_Y \ln Y + \mu(\mathbf{p})t + \mu_i t \tag{43}$$

where *Y* is output, $x = \sum_i p_i q_i$ is total expenditure on the inputs q_i, and as before $B(\mathbf{p}) > 0$ is homogeneous of degree one in prices and $\lambda(\mathbf{p}) \geq 0$ is homogeneous of degree zero in prices. There are two new elements here. First, the parameter β_Y measures overall economies of scale. When there are no input biases, ie $B(\mathbf{p}) = 1$ and $\lambda(\mathbf{p}) = 0$, then $\beta_Y = 0$ implies

[30] The parallel between cost and expenditure functions would be complete if individuals were able to learn over time how to make better use of goods and services in order to generate more utility. In some cases there is very suggestive evidence of a social learning process. The death toll before the Second World War on the roads in Great Britain peaked in 1938 when 6,648 people were killed, of whom 3,046 were pedestrians. By 2006 the annual death toll had fallen to 3,172, of whom 673 were pedestrians, and the death rate per capita had dropped to a third of the earlier level, even though the number of vehicles per capita increased to more than 8 times its 1938 level. (Source: *Annual Abstract*, various issues). Of course, many things changed over this period but one of them was surely that the habit of looking both ways before stepping into the road became more deeply engrained.

constant returns to scale and $\beta_Y < 0$ implies increasing returns. In this case the cost function is homothetic but not necessarily homogeneous of degree one in output. Second, the last two terms on the right hand side of (43) measure technical progress. Neutral technical progress is measured by the parameter μ_t ($\mu_t < 0$ implies that technical progress is positive); input-biased technical progress is measured by the function $\mu(\mathbf{p})$.[31] By analogy with $\lambda(\mathbf{p})$, $\mu(\mathbf{p})$ could be specified as

$$\mu(\mathbf{p}) = \sum_{i=1}^{i=N} \mu_i \ln p_i, \; \sum_{i=1}^{i=N} \mu_i = 0 \tag{44}$$

Under this specification, and with $B(\mathbf{p})$ and $\lambda(\mathbf{p})$ defined as earlier for the QAIDS (see (28) and (29)), the share equations are now given by:[32]

$$s_i = \frac{\partial \ln C}{\partial \ln p_i} = \frac{\partial \ln A(\mathbf{p})}{\partial \ln p_i} + \beta_i \left[\frac{B(\mathbf{p}) \ln Y}{1 - \lambda(\mathbf{p}) \ln Y} \right] + \frac{\lambda_i}{B(\mathbf{p})} \left[\frac{B(\mathbf{p}) \ln Y}{1 - \lambda(\mathbf{p}) \ln Y} \right]^2 + \mu_i t \tag{45}$$

The parameters β_i and λ_i now measure input bias in scale economies. If they are all zero there is no bias and the degree of returns to scale is measured just by β_Y. The parameter μ_i measures the bias in technical progress against input *i*: $\mu_i < 0$ would imply that technical progress is biased in favour of input *i*.

If our goal is to estimate the degree of economies of scale and the rate of technical progress, the parameters of interest in the cost function can be estimated by a simpler method than in the case of the expenditure function. We can just replace the price variables in (45) by principal components and then estimate the β_i, λ_i and μ_i, while imposing the appropriate cross-equation restrictions. Next, the degree of scale economies and the rate of neutral technical progress can be estimated by differentiating (43) totally with respect to time, using (44), applying Shephard's Lemma, and rearranging:

$$\begin{aligned} &\frac{d \ln x(t,t)}{dt} - \sum_{i=1}^{i-N} s_i(t,t) \left(\frac{d \ln p_i(t)}{dt} \right) - \sum_{i=1}^{i=N} \mu_i \ln p_i(t) \\ &- \left[\frac{B(\mathbf{p}(t))}{\left[1 - \lambda(\mathbf{p}(t) \ln Y(t)\right]^2} \right] \left(\frac{d \ln Y(t)}{dt} \right) = \mu_t + \beta_Y \left(\frac{d \ln Y(t)}{dt} \right) \end{aligned} \tag{46}$$

---

31 A cheap generalisation would be to add terms in $(\ln Y)^2$ and t^2 to the right hand side of equation (43).

32 These are cost shares, not revenue shares. In the presence of economies of scale there may be monopoly power, so profit is above the competitive level. I assume that the competitive rate of return to capital is known so that it is possible to calculate competitive rental prices for capital inputs (see Oulton (2007) for alternative ways of doing this).

Everything on the left hand side is now measurable and the only unknowns are the coefficients μ_t and β_Y on the right hand side. So (46) can be considered as a regression equation and used to estimate these remaining unknowns.[33]

The compensated shares, holding output constant at its level in period *b*, are

$$s_i(t,b) = \frac{\partial \ln A(\mathbf{p}(t))}{\partial \ln p_i(t)} + \beta_i \left[\frac{B(\mathbf{p}(t)) \ln Y(b)}{1 - \lambda(\mathbf{p}(t)) \ln Y(b)} \right] + \frac{\lambda_i}{B(\mathbf{p}(t))} \left[\frac{B(\mathbf{p}(t)) \ln Y(b)}{1 - \lambda(\mathbf{p}(t)) \ln Y(b)} \right]^2 + \mu_i t$$

$$= \frac{\partial \ln A(\mathbf{p}(t))}{\partial \ln p_i(t)} + \mu_i t \tag{47}$$

setting $\ln Y(b) = 0$. So the relationship between the actual and the compensated shares is

$$s_i(t,b) = s_i(t,t) - \beta_i \left[\frac{B(\mathbf{p}(t)) \ln Y(t)}{1 - \lambda(\mathbf{p}(t)) \ln Y(t)} \right] - \frac{\lambda_i}{B(\mathbf{p}(t))} \left[\frac{B(\mathbf{p}(t)) \ln Y(t)}{1 - \lambda(\mathbf{p}(t)) \ln Y(t)} \right]^2 \tag{48}$$

and the compensated shares can be used to construct a Konüs index of input prices.

The analysis of inequality in the preceding sub-section can also be applied to the cost functions of firms, if the size distribution varies over time or across countries. Entropy (*I*) and the related statistic *J* would now appear in the share equations (45), just as they do in (39).

Finally, an interesting question is whether anything useful can be concluded when output is not in fact measurable. In many private services, the inputs may be measured fairly easily but we don't know how to measure real output very well. This suggests that we might follow the same strategy as in the case of consumer demand. In that case, we eliminated unmeasured utility from the right hand side of the share equations by substituting from the expenditure function. The shares thus became functions of deflated expenditure (see equations (30)). Could the same strategy work for cost functions? Unfortunately not. If we rearrange the cost function (43) we obtain:

$$\frac{B(\mathbf{p}) \ln Y}{1 - \lambda(\mathbf{p}) \ln Y} = \ln \left[\frac{x}{A(\mathbf{p})} \right] - \left[\beta_Y \ln Y + \mu(\mathbf{p}) t + \mu_t t \right]$$

If we substitute this expression into the share equations (45) we are still left with the problem of estimating the unknown coefficients μ_t and β_Y and we still need a measure of real output.

---

33 Actually, overall technical progress is not separately identifiable from biased technical progress. Any non-zero estimate for μ_t can be absorbed into the μ_i by relaxing the constraint that $\sum_i \mu_i = 0$.

The root of the problem is that real output is necessarily cardinal while utility is only ordinal. And for utility there is no counterpart to technical progress.[34]

5. Conclusions

An algorithm which generates Konüs price indices when demand is not homothetic has now been presented. We have shown that it can be applied in both time series and cross-section. It is not dependent on the assumption of a representative consumer but can be extended to the case where income levels differ between consumers. The same algorithm can be applied to the parallel problem of estimating a true index of a producer's input prices. The algorithm involves some econometric estimation but uses exactly the same data, neither more nor less, as are required for conventional index numbers, namely prices and quantities.

It is now time to consider some limitations of the analysis and some unanswered questions. If we are trying to measure the standard of living, then our maintained hypothesis must be that tastes are identical. Otherwise the relative living standards of (say) Bangladeshi peasants and American investment bankers must be regarded as simply incommensurable. But the assumption of identical tastes might be considered overly strong. Is an intermediate position possible, in which tastes are identical at some comparatively high level, but might differ at a lower one? For example, the taste for hot, non-alcoholic beverages might be universal even though (at identical incomes and prices) some people prefer tea and others prefer coffee.

A related and unanswered question in the theory of demand and production is, at what level of aggregation is the analysis supposed to apply? It is hard to believe that there exists a stable structure of preferences (common to all time periods and all countries) at a very detailed level, such as individual brands of breakfast cereal. Equally, it is not obvious that "food" is the right level either, since food items range from necessities (bread) to luxuries (caviar). In practice, the level of aggregation is often chosen on pragmatic grounds, to obtain sufficient observations to estimate the parameters of interest.

---

[34] In special cases the problem is solvable. Mellander (1992) shows that we can deduce real output in the case where input demand is homothetic, there are decreasing returns to scale, and the mark-up of price over marginal cost is constant. Then the ratio of the value of output (assumed observable) to the value of total cost is an indicator of the degree of returns to scale.

Finally, the index numbers developed here are only “true” if the underlying theory is correct and also applicable to the problem at hand. The economic theory applied in this paper has been static. It is likely that agents’ choices include an intertemporal element: in deciding whether or not to purchase a line of cocaine, the consumer may consider the future consequences as well as current income and relative prices. Habit may be important even in the absence of addiction as the macro literature has emphasised. If so, index numbers should reflect an intertemporal element too.

ANNEX

A.1 The Konüs price index lies between the compensated Laspeyres and the compensated Paasche indices

This proposition follows from the well known inequalities for the Laspeyres and Paasche derived by Konüs (1939); see also Deaton and Muellbauer (1980), chapter 7. By definition of the expenditure function, we have:

$$\sum_{i=1}^{i=N} p_i(t)q_i(t-1) \geq E[\mathbf{p}(t), u(t-1)]$$

Denoting the Laspeyres price index for year t with base year $t-1$ by $P^L(t,t-1)$, it follows that

$$P^L(t,t-1) = \frac{\sum_{i=1}^{i=N} p_i(t)q_i(t-1)}{\sum_{i=1}^{i=N} p_i(t-1)q_i(t-1)} \geq \frac{E[\mathbf{p}(t), u(t-1)]}{E[\mathbf{p}(t-1), u(t-1)]}$$

since $\sum_i p_i(t-1)q_i(t-1) = E[\mathbf{p}(t-1), u(t-1)]$. By definition of the expenditure function again,

$$\sum_{i=1}^{i=N} p_i(t-1)q_i(t) \geq E[\mathbf{p}(t-1), u(t)]$$

whence

$$P^P(t,t-1) = \frac{\sum_{i=1}^{i=N} p_i(t)q_i(t)}{\sum_{i=1}^{i=N} p_i(t-1)q_i(t)} \leq \frac{E[\mathbf{p}(t), u(t)]}{E[\mathbf{p}(t-1), u(t)]}$$

where $P^P(t,t-1)$ is the Paasche price index for year t with base year $t-1$. Now in the present case utility is being held constant at the level of period b, ie $u(t-1) = u(t) = u(b)$, so we have

$$P^L(t,t-1) \geq P^K(t,b)/P^K(t-1,b) \geq P^P(t,t-1) \qquad \text{(A1)}$$

where $P^K(t,b) = \dfrac{E[\mathbf{p}(t), u(b)]}{E[\mathbf{p}(b), u(b)]}$ is the Konüs price index with base year b. In recognition of the fact that this proposition holds when utility is held constant at the level of period b, in the text we refer to these Laspeyres and Paasche indices as compensated ones and write them as $P^L(t,t-1,b)$ and $P^P(t,t-1,b)$ respectively.

Since the compensated Fisher index is the geometric mean of the compensated Laspeyres and the compensated Paasche, like the Konüs it must always lie between the Laspeyres and the Paasche:

$$P^L(t,t-1) \geq P^F(t,b)/P^F(t-1,b) \geq P^P(t,t-1) \tag{A2}$$

A.2 Aggregating over unequal incomes in the QAIDS

As given in equation (38), repeated here for convenience, the share of product i in aggregate expenditure is a weighted average of the shares of the various income groups:

$$\begin{aligned} s_i = \sum_{g=1}^{g=G} w_g s_{ig} &= \alpha_i + \sum_{j=1}^{j=N} \gamma_{ij} \ln p_j + \beta_i \sum_{g=1}^{g=G} w_g \ln x_g - \beta_i \ln A(\mathbf{p}) \\ &+ \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}} \left[\sum_{g=1}^{g=G} w_g (\ln x_g)^2 - 2\ln A(\mathbf{p}) \sum_{g=1}^{g=G} w_g \ln x_g + [\ln A(\mathbf{p})]^2 \right] \end{aligned} \tag{A3}$$

From (37),

$$\sum_{g=1}^{g=G} w_g \ln x_g = \sum_{g=1}^{g=G} w_g \ln(w_g G x) = \sum_{g=1}^{g=G} w_g \ln w_g + \ln G + \ln x$$

Also, from (37) again,

$$\begin{aligned} \sum_{g=1}^{g=G} w_g (\ln x_g)^2 &= \sum_{g=1}^{g=G} w_g (\ln w_g + \ln(Gx))^2 \\ &= \sum_{g=1}^{g=G} w_g \left[(\ln w_g)^2 + 2(\ln G + \ln x) \ln w_g + (\ln G + \ln x)^2 \right] \\ &= \sum_{g=1}^{g=G} w_g (\ln w_g)^2 + 2\ln G \sum_{g=1}^{g=G} w_g \ln w_g + 2\ln x \sum_{g=1}^{g=G} w_g \ln w_g \\ &\quad + \left[(\ln G)^2 + 2\ln G \ln x + (\ln x)^2 \right] \sum_{g=1}^{g=G} w_g \\ &= \sum_{g=1}^{g=G} w_g (\ln w_g)^2 + 2\ln G \sum_{g=1}^{g=G} w_g \ln w_g + 2\ln x \left[\sum_{g=1}^{g=G} w_g \ln w_g + \ln G \right] \\ &\quad + (\ln G)^2 + (\ln x)^2 \end{aligned} \tag{A4}$$

Therefore, plugging these results into (A3):

$$s_i=\sum_{g=1}^{g=G} w_g s_{ig}=\alpha_i+\sum_{j=1}^{j=N}\gamma_{ij}\ln p_j+\beta_i\sum_{g=1}^{g=G} w_g \ln x_g-\beta_i \ln A(\mathbf{p})$$

$$+\frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left[\sum_{g=1}^{g=G} w_g(\ln x_g)^2-2\ln A(\mathbf{p})\sum_{g=1}^{g=G} w_g \ln x_g+[\ln A(\mathbf{p})]^2\right]$$

$$=[\alpha_i+\beta_i\ln G]+\beta_i\sum_{g=1}^{g=G} w_g \ln w_g+\sum_{j=1}^{j=N}\gamma_{ij}\ln p_j+\beta_i\ln\left[\frac{x}{A(\mathbf{p})}\right]$$

$$+\frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left[(\ln x)^2-2\ln A(\mathbf{p})\ln(x)+[\ln A(\mathbf{p})]^2\right]$$

$$+\ \frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left[\sum_{g=1}^{g=G} w_g(\ln w_g)^2+2\ln G\sum_{g=1}^{g=G} w_g \ln w_g+2\ln x\sum_{g=1}^{g=G} w_g \ln w_g\right.$$

$$\left.+(\ln G)^2+2\ln x\ln G-2\ln A(\mathbf{p})\left[\sum_{g=1}^{g=G} w_g \ln w_g+\ln G\right]\right]$$

Therefore

$$s_i=[\alpha_i+\beta_i\ln G]+\sum_{j=1}^{j=N}\gamma_{ij}\ln p_j$$

$$-\beta_i I+\frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left[J-2I\ln G+(\ln G)^2\right]$$

$$+\left(\beta_i+\frac{2\lambda_i(\ln G-I)}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\right)\ln\left[\frac{x}{A(\mathbf{p})}\right]+\frac{\lambda_i}{\prod_{k=1}^{k=N} p_k^{\beta_k}}\left\{\ln\left[\frac{x}{A(\mathbf{p})}\right]\right\}^2$$

where we have set $I=-\sum_{g=1}^{g-G} w_g \ln w_g$ and $J=\sum_{g=1}^{g-G} w_g(\ln w_g)^2$. This is equation (39) of the main text.

A.3 The effects on the PIGLOG expenditure function of normalising deflated expenditure to equal one in the base year

I consider here in more detail the effect on the PIGLOG expenditure function $\ln x=\ln A(\mathbf{p})+B(\mathbf{p})\ln u$ of adopting the normalisation $x_b(b,b)=A(\mathbf{p}(b))$. The subscript b is intended to indicate that this normalisation is different for every possible choice of base year. In the main text this subscript is omitted to simplify the notation.

First note a general point about the choice of quantity units for the products. Suppose that initially the quantity unit for sugar is the kilo and the price per kilo of sugar is \$1. Then the quantity unit is changed to 2 kilos. Now the price per 2 kilos is \$2. But the quantity purchased of 2 kilo units is half that of 1 kilo units. So total expenditure on sugar is unchanged. (NB: we are considering here changes that an analyst might make; the quantity unit for sugar in a real shop is not being changed — it might be 1 kilo, 2 kilos or something else). Setting all prices equal to one in a reference year amounts to changing the quantity unit for each product. Eg if the quantity unit for the *i*th product was originally one kilo it is now the number of kilos which could have been purchased for $p'(r)$ monetary units in the reference year *r*, where the prime denotes the original units; after normalising all prices to equal one in the reference year, the price of the *i*th product in the new quantity units is $p_i(t) = p_i'(t) / p_i'(r)$. This normalisation leaves the value of total expenditure unchanged: changes in the prices are accompanied by offsetting changes in the quantities:

$$x(t,t) = \sum_i p_i(t) q_i(t) = \sum_i [p_i'(t) / p_i'(r)][p_i'(r) q_i'(t)] = \sum_i [p_i'(t) q_i'(t)]$$

Following this prior normalisation of prices, suppose that $\ln A(\mathbf{p}(b)) = h_b + \ln x(b,b)$. Then define a new monetary unit such that

$$\ln x_b(b,b) = h_b + \ln x(b,b) = \ln A(\mathbf{p}(b))$$

That is, putting $H_b = \exp(h_b)$,

$$x_b(b,b) = \sum_i H_b p_i q_i = \sum_i [H_b p_i'(b) / H_b p_i'(r)][H_b p_i'(r) q_i(b)]$$

$$= \sum_i [p_i'(b) / p_i'(r)][H_b p_i'(r) q_i(b)]$$

(This would be like changing the monetary unit from dollars to cents). Then in the new monetary unit we also have

$$\ln x_b(t,b) = h_b + \ln x(t,b) \tag{A.5}$$

If prices have already been normalised, then changing the monetary unit leaves the p_i and therefore the value of $B(\mathbf{p})$ unaffected (in the case of the AIDS the β_i are semi-elasticities of the shares with respect to changes in *x* and so are unaffected by any changes in the monetary unit). The expenditure function at time *b* can now be written as:

$$\ln x_b(b,b) = \ln A(\mathbf{p}(b)) + B(\mathbf{p}(b)) \ln u_b(b) \tag{A.6}$$

whence in the new units

$$\ln u_b(b) = 0 \tag{A.7}$$

Here a subscript b has been added to utility to indicate that changing the monetary unit involves changing the utility unit as well: in fact, $u_b(t) = u(t)/u(b)$: see below. The expenditure function at time t with the utility level of time b is now:

$$\ln x_b(t,b) = \ln A_b(\mathbf{p}(t)) + B(\mathbf{p}(t)) \ln u_b(b)$$

whence, using (A.7),

$$\ln x_b(t,b) = \ln A_b(\mathbf{p}(t)) \tag{A.8}$$

Note that I have added a subscript b to the $A(\mathbf{p}(t))$ function since the normalisation generally changes this. This is the case even though the *value* of the function is by definition the same at time b as before the normalisation, ie $A_b(\mathbf{p}(b)) = A(\mathbf{p}(b))$: see below.

Now consider the effect of the normalisation on $A(\mathbf{p})$. With the original monetary unit, we have:

$$\ln x(t,b) = \ln A(\mathbf{p}(t)) + B(\mathbf{p}(t)) \ln u(b) \tag{A.9}$$

In the new units, noting that $h_b = B(\mathbf{p}(b) \ln u(b)$,

$$\ln x_b(t,b) = x(t,b) - h_b = \ln A(\mathbf{p}(t)) + B(\mathbf{p}(t)) \ln u(t) - h_b$$

$$= \ln A(\mathbf{p}(t)) + \left[B(\mathbf{p}(t)) - B(\mathbf{p}(b))\right] \ln u(b) + B(\mathbf{p}(t))\left[\ln u(t) - \ln u(b)\right] \tag{A.10}$$

$$= \ln A_b(\mathbf{p}(t)) + B(\mathbf{p}(t)) \ln u_b(t)$$

That is,

$$\ln A_b(\mathbf{p}(t)) = \ln A(\mathbf{p}(t)) + \left[B(\mathbf{p}(t)) - B(\mathbf{p}(b))\right] \ln u(b) \tag{A.11}$$

and

$$\ln u_b(t) = \ln u(t) - \ln u(b) \tag{A.12}$$

Finally, note that from (A.11) $A_b(\mathbf{p}(b)) = A(\mathbf{p}(b))$, as asserted above.

References

Anand, S. and P. Segal (2008). "What do we know about global income inequality?". *Journal of Economic Literature*, 46:1, 57-94.

Apostol, T. (1957). *Mathematical Analysis*. Addison-Wesley: Reading, Mass.

Balk, B.M. (1995). "Axiomatic price index theory: a survey." *International Statistical Review*, 63 (1), 69-93.

Balk, B.M. (2004). "Direct and chained indices: a review of two paradigms." Paper presented to the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver, 2004; accessed on 17 December 2007 at http://www.ipeer.ca/papers/Balk,%20Sept.15,2004,%20Vancouver%20rev.pdf.

Balk, B.M. (2005). "Divisia price and quantity indices: 80 years after." *Statistica Neerlandica*, 59, 119-158.

Banks, J., R. Blundell and A. Lewbel (1997). "Quadratic Engel curves and consumer demand." *The Review of Economics and Statistics*, 79(4), 527-539.

Blow, L., A. Leicester, and Z. Oldfield (2004). *Consumption Trends in the UK, 1975-99*. The Institute for Fiscal Studies, London.

Blundell, R., X. Chen and D. Christensen (2007). "Semi-nonparametric IV estimation of shape-invariant Engel curves". *Econometrica*, 75(6), 1613-1669.

Caves, D.W., L.R. Christensen, and W.E. Diewert (1982). "Multilateral comparisons of output, input, and productivity using superlative index numbers." *Economic Journal*, 92, 73-86.

Christensen, L.R., D.W. Jorgenson and L.J. Lau (1975). "Transcendental logarithmic utility functions". *American Economic Review*, 65 (June), 367-383.

Deaton, A. and J. Muellbauer (1980a). "An almost ideal demand system." *American Economic Review*, 70, 312-336.

Deaton, A. and J. Muellbauer (1980b). *Economics and Consumer Behaviour*. Cambridge University Press: Cambridge.

Diewert, W.E. (1976). "Exact and superlative index numbers." *Journal of Econometrics*, 4, 115-46.

Diewert, W.E. (1981). "The economic theory of index numbers: a survey", in: W.E. Diewert and A.O. Nakamura, (eds.), *Essays in Index Number Theory Volume 1*, 1993. North-Holland: Amsterdam.

Diewert, W.E. (2002). "Similarity and dissimilarity indexes: an axiomatic approach". Department of Economics, University of British Columbia, no. 10-02.

Diewert, W.E. (2008). "Index numbers". In *The New Palgrave Dictionary of Economics*, edited by Steven N. Durlauf and Lawrence Blume. Palgrave Macmillan.

Diewert, W.E. (2009). "Cost of living indexes and exact index numbers". In *Quantifying Consumer Preferences*, edited by Daniel Slottje. Emerald Publishing Group.

Diewert, W.E. and T.J. Wales (1988). "A normalized quadratic semiflexible functional form". *Journal of Econometrics*, 37, 327-342.

Divisia, F. (1925-1926). "L'indice monétaire et la théorie de la monnaie." *Revue d'Economie Politique*, 39(4), 842-864; 39(5), 980-1008; 39(6), 1121-1151 [1925]; 40(1), 49-81 [1926].

Dowrick, S. and J. Quiggin, J. (1997). "True measures of GDP and convergence". *American Economic Review*, 87(1), March, 41-64.

Engel, E. (1857) [1895] "Die productions- und consumptionverhältnisse des Königsreich Sachsen". Reprinted in *Bulletin de l'Institut International de Statistique*, IX.

Feenstra, R.C. and M.B. Reinsdorf (2000) "An exact price index for the almost ideal demand system." *Economics Letters*, 66, 159-162.

Gorman, W.M. (1981). “Some Engel curves”. In *Essays in the Theory and Measurement of Consumer Behaviour: In Honour of Richard Stone*, edited by Angus Deaton, 7-30. Cambridge: Cambridge University Press.

Hill, R.J. (1997). “A taxonomy of multilateral methods for making international comparisons of prices and quantities.” *The Review of Income and Wealth*, 43(1, March), 49-69.

Hill, R.J. (1999). “Comparing price levels across countries using minimum spanning trees”. *Review of Economics and Statistics*, 81(1), 135-142.

Hill, R.J. (2004). “Constructing price indexes across space and time: the case of the European Union.” *American Economic Review*, 94 (5, December), 1379-1410.

Hill, R.J. (2006). “Superlative index numbers: not all of them are super”. *Journal of Econometrics*, 130, 25-43.

Houthakker, H.S. (1957). “An international comparisons of household expenditure patterns commemorating the centenary of Engel’s Law”. *Econometrica*, 25, 532-551.

Hulten, C.R. (1973). “Divisia index numbers.” *Econometrica*, 41, 1017-1025.

Johnson, R.A. and D.W. Wichern (2002). *Applied Multivariate Statistical Analysis*. Prentice-Hall Inc, New Jersey.

Konüs, A.A. (1939). “The problem of the true index of the cost-of-living.” *Econometrica*, 7, 10-29. (English translation; first published in Russian in 1924).

Lewbel, A. (1991). “The rank of demand systems: theory and nonparametric estimation”. *Econometrica*, 59(3), 711-730.

Lewbel, A. and K. Pendakur (2009). “Tricks with Hicks: the EASI demand system”. *American Economic Review*, 99:3 (June), 827-863.

Marshall, A. (1920) [1966]. *Principles of Economics* (8^{th} edition). Macmillan: London.

Mellander, E. (1992). “An indirect approach to measuring productivity in private services”. *Scandinavian Journal of Economics*, 94, Supplement, 229-244.

Muellbauer, J. (1976). “Community preferences and the representative consumer.” *Econometrica*, 44, 524-543.

Neary, J.P. (2004). “Rationalizing the Penn World Table: true multilateral indices for international comparisons of real income.” *American Economic Review*, 94 (5),1411-1428.

Office for National Statistics (1998). *The Retail Prices Index: Technical Manual: 1998 Edition*. The Stationery Office, London. Available at www.statistics.gov.uk.

Office for National Statistics (2006). *Consumer Price Indices: Technical Manual (2006 Edition)*. Office for National Statistics, London. Available at www.statistics.gov.uk.

Oulton, N. (2007). “Ex post versus ante ex measures of the user cost of capital”. *Review of Income and Wealth*, Series 53, No. 2 (June), 295-317.

Oulton, N. (2008). “Chain indices of the cost of living and the path-dependence problem: an empirical solution.” *Journal of Econometrics*, 144: 1 (May), 306-324. [http://dx.doi.org/10.1016/j.jeconom.2008.02.001; discussion paper version: http://cep.lse.ac.uk/pubs/download/dp0797.pdf].

Samuelson, P.A. and S. Swamy (1974). “Invariant economic index numbers and canonical duality: survey and synthesis.” *American Economic Review*, 64, 566-593.

Stigler, G.J. and G.S. Becker (1977). “De gustibus non est disputandum”. *American Economic Review*, 67(2), March, 76-90; stable URL: http://www.jstor.org/stable/1807222.

Theil, H. (1967). *Economics and information theory*. North-Holland: Amsterdam.

Tobin, J. (1987). “Fisher, Irving”. In *The New Palgrave: A Dictionary of Economics*, edited by John Eatwell, Murray Milgate, and Peter Newman. Palgrave: Basingstoke. Reprinted in *The New Palgrave Dictionary of Economics*, edited by Steven N. Durlauf and Lawrence Blume, 2008. Palgrave Macmillan.

U.S. Bureau of Labor Statistics (2007). "The Consumer Price Index (updated 06/2007)". Chapter 17 of *BLS Handbook of Methods*. Accessed at http://www.bls.gov/opub/hom/pdf/homch17.pdf on 4 June 2009.

World Bank (2008). *Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program*. Available at http://siteresources.worldbank.org/ICPINT/Resources/icp-final.pdf.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

948	Alex Bryson Bernd Frick Rob Simmons	The Returns to Scarce Talent: Footedness and Player Remuneration in European Soccer
947	Jonathan Wadsworth	Did the National Minimum Wage Affect UK Wages?
946	David Marsden	The Paradox of Performance Related Pay Systems: 'Why Do We Keep Adopting Them in the Face of Evidence that they Fail to Motivate?'
945	David Marsden Almudena Cañibano	Participation in Organisations: Economic Approaches
944	Andreas Georgiadis Alan Manning	One Nation Under a Groove? Identity and Multiculturalism in Britain
943	Andreas Georgiadis Alan Manning	Theory of Values
942	Kristian Behrens Giordano Mion Yasusada Murata Jens Südekum	Trade, Wages and Productivity
941	David Marsden Richard Belfield	Institutions and the Management of Human Resources: Incentive Pay Systems in France and Great Britain
940	Elhanan Helpman Oleg Itskhoki Stephen Redding	Inequality and Unemployment in a Global Economy
939	Norman Ireland Robin A. Naylor Jeremy Smith Shqiponja Telhaj	Educational Returns, Ability Composition and Cohort Effects: Theory and Evidence for Cohorts of Early-Career UK Graduates
938	Guy Mayraz Jürgen Schupp Gert Wagner	Life Satisfaction and Relative Income: Perceptions and Evidence
937	Nicholas Bloom Raffaella Sadun John Van Reenen	The Organization of Firms Across Countries
936	Jean-Baptiste Michau	Unemployment Insurance and Cultural Transmission: Theory and Application to European Unemployment

935	João M. C. Santos-Silva Silvana Tenreyro	Trading Partners and Trading Volumes: Implementing the Helpman-Melitz-Rubinstein Model Empirically
934	Christian Morrisson Fabrice Murtin	The Century of Education
933	João M. C. Santos-Silva Silvana Tenreyro	Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator
932	João M. C. Santos-Silva Silvana Tenreyro	On the Existence of the Maximum Likelihood Estimates for Poisson Regressioin
931	Richard Freeman John Van Reenen	What If Congress Doubled R&D Spending on the Physical Sciences?
930	Hector Calvo-Pardo Caroline Freund Emanuel Ornelas	The ASEAN Free Trade Agreement: Impact on Trade Flows and External Trade Barriers
929	Dan Anderberg Arnaud Chevalier Jonathan Wadsworth	Anatomy of a Health Scare: Education, Income and the MMR Controversy in the UK
928	Christos Genakos Mario Pagliero	Risk Taking and Performance in Multistage Tournaments: Evidence from Weightlifting Competitions
927	Nick Bloom Luis Garicano Raffaella Sadun John Van Reenen	The Distinct Effects of Information Technology and Communication Technology on Firm Organization
926	Reyn van Ewijk	Long-term health effects on the next generation of Ramadan fasting during pregnancy
925	Stephen J. Redding	The Empirics of New Economic Geography
924	Rafael Gomez Alex Bryson Tobias Kretschmer Paul Willman	Employee Voice and Private Sector Workplace Outcomes in Britain, 1980-2004
923	Bianca De Paoli	Monetary Policy Under Alterative Asset Market Structures: the Case of a Small Open Economy
922	L. Rachel Ngai Silvana Tenreyro	Hot and Cold Seasons in the Housing Market

The Centre for Economic Performance Publications Unit
Tel 020 7955 7284 Fax 020 7955 7595 Email info@cep.lse.ac.uk
Web site http://cep.lse.ac.uk