FISCAL STUDIES, vol. 26, no. 2, pp. 197–224 (2005) 0143-5671

The Elasticity of Marginal Utility of Consumption: Estimates for 20 OECD Countries*

DAVID J. EVANS†

†*Business School, Oxford Brookes University*
(djevans@brookes.ac.uk)

Abstract

In social project appraisal, the policy profile of both distributional welfare weights and the social discount rate has risen considerably in recent years. This fact has important implications for the allocation of funds to social projects and policies in countries, and in unions of countries such as the EU. A key component in the formulae for both welfare weights and the social discount rate is the elasticity of marginal utility of consumption, *e*. A critical review of existing evidence on *e* suggests that the UK Treasury's preferred value of unity is too low. New evidence presented in this paper, based on the structure of personal income tax rates, suggests that, on average, for developed countries *e* is close to 1.4. This particular approach to the estimation of *e* has previously been under-utilised by researchers.

I. Introduction

The elasticity of marginal utility of consumption, *e*, is an important concern in welfare economics. In a policy context, its profile has been raised recently as a result of revised guidance from the UK Treasury on the appraisal and evaluation of public sector projects and policies (see HM Treasury (2003)). There are two key areas in which the policy status of *e* has been raised – deliberations concerning an appropriate social discount rate and government

*The author would like to thank the anonymous referees for their helpful comments and suggestions for revising the original draft.
JEL classification numbers: D60, D61, H24, R13.

emphasis on distributional welfare weights (HM Treasury, 2003, annex 6, pages 97–100 and annex 5, pages 91–96). The new Treasury guidance emerged from extensive consultation with both practitioners and academics. For example, the work of Pearce and Ulph (1999), Cowell and Gardiner (1999) and OXERA (2002) has been especially influential.

The new policy initiatives in the UK concerning both distributional welfare weights and discount rates have important implications for policy practice in other countries. In the interests of a consistent approach to the measurement of discount rates and welfare weights, other countries should at least consider adopting the same broad procedures in the cost–benefit analysis of social projects and policies. This concern over consistency of policy approach applies especially to member countries of the European Union.

1. The social discount rate and *e*

The fundamental importance of *e* in relation to the social discount rate has been considerably enhanced as a result of the UK Treasury's decision to focus entirely on the social time preference rate (stpr) (see HM Treasury (2003)). This contrasts with previous practice, whereby the discount rate reflected cost-of-capital considerations to an important extent (see HM Treasury (1997)). Of course, *e* is now a specific component element of the social discount rate, as the following stpr formula (Ramsey, 1928) reveals:

(1) $$stpr = p + eg$$

where p is the utility discount rate, e the elasticity of marginal utility of consumption (absolute value) and g the projected long-run average annual rate of growth in per-capita real consumption.

In fact, from the work of Ramsey (1928), it is clear that marginal utility and the related concept of *e* have other important applications. For example, he used the marginal utility of money to find an answer to the question 'How much of its income should a nation save?'. Also, the social discount rate has an important application in the economic analysis of the extraction of exhaustible resources. See, for example, Hotelling (1931) and Slade (1982), who argue that for extraction of non-renewable resources to be justified, the net social price must go up in line with the social discount rate. It can be seen from equation (1) that an estimate of *e* is required in order to calculate an appropriate rate of discount.

2. Distributional welfare weights and *e*

In the estimation of distributional welfare weights for application in the appraisal of social projects and policies impacting on different socio-economic groups, *e* also has a central role. The size of *e* determines the extent to which marginal social utility declines as income rises, and knowledge of its value permits a comparison of relative marginal utilities for groups with contrasting per-capita real incomes. In a regional context, for example, the appropriate welfare weight expression is given in equation (2) below. Welfare weights could be used by governments in reaching decisions on the allocation of funds to social projects and policies on a regional basis.

$$(2) \qquad W_i = \left(\frac{C^*}{C_i} \right)^e$$

where W_i is the distributional welfare weight for region i, C_i is per-capita real consumption in region i and C^* is per-capita real consumption in the country as a whole (or union of countries, e.g. the EU). For measures of regional welfare weights in the cases of India and the UK respectively, see Kula (2002) and Evans, Kula and Sezer (2005).

3. The importance of *e* and a need to review existing empirical evidence

For the purpose of estimating appropriate values of both distributional welfare weights and social discount rates for countries, the ‘correct’ measurements of *e* are a crucial policy concern. There have been many attempts to estimate *e*, but very different approaches have been used, often resulting in widely differing values. So, in Section II of the paper, the more recent empirical evidence on *e* will be critically reviewed, especially those studies that have contributed significantly to the decision of the UK Treasury to reduce its preferred value of *e* from 1.5 to 1.0. Such a sizeable reduction has potentially serious implications for the allocation of funds to social projects, both on an intergenerational basis and between different socio-economic groups.

It is argued that undue emphasis has been placed on one particular method of estimating *e* based on lifetime consumption behaviour. Other estimation approaches are worthy of some respect – for example, the Fellner (1967) model based on the demand for want-independent consumer goods. Some limited evidence from a tax-based approach relating to the revealed social values of governments is also given due consideration. Both Stern (1977) and Cowell and Gardiner (1999) provided estimates of *e* based on personal income tax data for the UK and the results suggested an *e* value well in excess of unity. There is much more scope for work on tax-based methods of estimating *e*, especially in relation to countries other than the UK.

4. The need for new estimates of *e*

From the critical review of the empirical evidence on *e* emerges a need for more work on the estimation of this important parameter, work covering a wide range of countries. In Section III of the paper, the results from new research are presented – *e* estimates based on personal income tax rates for 20 OECD countries using the 'equal absolute sacrifice model'. Tax data relating to wage earnings in a recent year are used. The empirical work involves formal statistical analysis and employs both parametric and non-parametric tests. The main results suggest that an average measure of *e* for countries is close to 1.4. Furthermore, statistical tests reveal that the tax data, taken at contrasting levels of wage income, offer support for an iso-elastic social utility function. Using a different empirical approach, Blue and Tweeten (1997) also found evidence for constancy of *e* in the case of the USA.

5. Policy implications and overall evidence on *e*

In Section IV of the paper, social time preference rates incorporating tax-based estimates of *e* are presented for selected OECD countries and then briefly discussed. Also, some speculation is entertained concerning the use of welfare weights in a 'future Europe' for the purposes of determining both regional funding allocations for social projects and member countries' net contributions to the EU budget. This is a 'Europe' in which tax harmonisation and full monetary union have been realised. Finally, in the concluding section of the paper, a summary of the main findings concerning both the magnitudes and consistency of *e* values is reported. In the light of the results, the need for additional empirical work is also highlighted.

II. A review of recent empirical evidence on *e*

In selecting an *e* value of unity in its latest guidance on appraisal and evaluation in central government, HM Treasury draws heavily on the work of Cowell and Gardiner (1999), especially, and Pearce and Ulph (1995). For specific evidence of this fact, see HM Treasury (2003, annex 5, page 93).

The empirical work on *e* involves three fundamentally different approaches: direct survey methods, indirect behavioural evidence and revealed social values. The results for *e*, based on each approach, will be considered critically, focusing on the difficulties and problems involved. It will be argued that undue emphasis has been placed on just one particular method involving the analysis of lifetime consumption behaviour. A case is made for placing at least some emphasis on the Fellner demand model (see Fellner (1967)) plus a greater focus on a tax-based method relating to the revealed social values of governments. The main picture to emerge from this review of recent empirical evidence is that the results do not, on balance, support an *e* value as low as unity. However, further empirical work is still required for the purpose of obtaining a more reliable estimate. The nature of this additional work is given due emphasis.

1. Survey methods

Estimates of *e* based on survey evidence concerning both risk and inequality aversion are sensitive to the nature of the questions asked and the types of respondents targeted. For example, Amiel, Creedy and Hurn (1999) focused on students, so the low *e* values obtained (see Table 1) only apply to this particular group, one that is not representative of the general public. Barsky et al. (1995) focused on middle-aged persons in the USA, so the high *e* value of 4 does not necessarily hold for their UK counterparts, never mind the general public.

These surveys, concerned with the measurement of risk and inequality aversion, only elicit responses to questions relating to hypothetical circumstances. This important fact clearly raises the possibility of macho-posturing behaviour exhibited through the under-revealing of aversion, which is more likely the case with young persons such as students.

TABLE 1

Summary of e *measures based on different approaches*

		Estimates of e
Survey methods		
Amiel et al., 1999	Student surveys; inequality aversion	0.2–0.8
Barsky et al., 1995	US middle-aged; risk aversion	Approx. 4.0
Indirect behavioural evidence		
Blundell et al., 1994	1970–86 FES data (UK); lifetime consumption behaviour	1.20–1.40[a] 0.35–1.05[b]
Evans and Sezer, 2002	1967–97 (UK), FFF model; demand for food	1.60
Revealed social values		
Cowell and Gardiner, 1999	UK income tax 1999–2000:	
	Income tax only	1.41
	Tax + NICs	1.28
Evans and Sezer, 2004	UK income tax 2001–02	1.50

[a]Basic model of the Euler equation for consumption.
[b]A dummy variable is added to the basic model in order to capture the effect of high real interest rates in the 1980s.

Surveys are costly in terms of both time and money, and there is a need to cover more representative cross-sections of the general public in order to obtain appropriate *e* values. Even then, follow-up work should be seen to yield consistent results. Evidence from alternative approaches is clearly required.

2. Indirect behavioural evidence

Two contrasting approaches and the associated estimates of *e* will be considered: models of lifetime consumption behaviour and consumer demand models for preference-independent goods. The former approach has been especially influential in the UK Treasury's decision to set a unitary value for *e* and therefore requires close consideration (see HM Treasury (2003)).

(a) Lifetime consumption behaviour

The interpretation of *e* as the reciprocal of the intertemporal elasticity of substitution in consumption has found favour theoretically in recent times. The main reason for this is that it provides a measure of relative risk aversion (see, for example, Cowell and Gardiner (1999, pages 28–29)), so the objection to an *e* measure based on a utility-under-certainty approach – namely, that utility is an ordinal concept unless strong conditions are imposed on the utility function – does not apply in this intertemporal context. Both Cowell and Gardiner (1999) and Pearce and Ulph (1995) place considerable emphasis on this method and set out the relevant theory and the underlying algebra involved. Based on models of lifetime consumption behaviour, the major associated empirical studies have been conducted by Blundell, Browning and Meghir (1994), Attanasio and Browning (1995) and Besley and Meghir (1998). These studies cover different countries and mostly produce estimates of *e* that are close to unity, although normally just above this value. The study by Blundell et al. has been enormously influential at policy level in the UK and therefore their results merit special attention.

Results for *e* estimated by Blundell et al. (1994) are presented for both basic and adjusted models in Table 1; there is a striking contrast between them. The data used consist of a pooled sample of cross-section and time-series observations relating to consumption, income and the rate of interest, along with a range of socio-economic dummy variables. Very large samples of household data were taken from the Family Expenditure Survey (FES) over the period 1970–86.

In the basic model, *e* estimates range from 1.2 at the first decile income level for households with average family characteristics, to 1.4 at the ninth decile income level for corresponding households. This very modest rise in *e* with income level is not inconsistent with the approximate relevance of an iso-elastic social utility function. However, in the adjusted model, which includes a dummy variable to capture the effect of high UK real rates of interest in the 1980s, very much lower estimates of *e* are derived. As shown in Table 1, these estimates rise steeply with income, ranging between the values 0.35 and 1.05 at the first and ninth decile average family income levels. These latter results stem from a relatively crude dummy-variable adjustment to an otherwise sophisticated empirical model. In the judgement of this author, the results from the basic model do seem far more plausible and, as inspection of Table 1 reveals, are reasonably consistent with *e* estimates derived from other non-survey empirical methods.

While the two models used by Blundell et al. (1994) yield sharply contrasting estimates of *e*, it is interesting to note that if the mid-range estimates are averaged then a value of 1 is obtained. This just happens to be the same as the unitary *e* value assumed by the UK Treasury. Furthermore, since the Treasury draws heavily on the work of Cowell and Gardiner (1999), who, in turn, highlight the Blundell et al. study, this assumption is no coincidence. The assumption does not, however, sit easily with the evidence on *e* obtained from the application of other methods (see Table 1).

There are some problems in using estimates of *e* based on the work of Blundell et al. (1994) for the purpose of estimating both current values of UK distributional welfare weights and the social discount rate. These problems will be briefly discussed before considering the evidence on *e* associated with alternative estimation approaches.

Apart from which particular interest rate to use, the large difference between retail loan and saving rates needs to be taken properly into account, especially in the absence of fixed mark-ups by financial institutions. Furthermore, the data period 1970–86 finished nearly 20 years ago and the financial market environment was very different from that of ‘today’. The market was regulated and less competitive, and building societies, for example, were unable to compete as retail banks until the passing of the Building Societies Act (1986) at the end of the sample data period. The period 1970–86 was characterised by many market shocks and macroeconomic policy changes. These included major increases in oil prices, the introduction of flexible exchange rates, UK membership of the EU, record inflation levels and a major switch from fiscal policy to monetarism. Such market and policy turbulence makes it difficult to place much confidence in estimates of *e* derived from models of lifetime consumption behaviour over the period 1970–86. For the UK, estimation of these same models over the period 1987–2003 – a period of deregulated financial markets and relatively low inflation – should be regarded as a clear policy priority in relation to obtaining a more reliable measure of *e*.

(b) Consumer demand for a preference-independent good

This is an older method of estimating *e* which is based on the work of Fisher (1927), Frisch (1932) and Fellner (1967). It is commonly referred to as the FFF model. For a preference-independent good with a relatively small share in consumer budgets, it is demonstrated, to good approximation, that the value of *e* in this FFF model is given by the ratio of the income elasticity of demand to the compensated own-price elasticity. For products with more significant budget shares, an adjustment to this ratio of elasticities is required: either the application of the Frisch formula (see Frisch (1959)) or the Amundsen modification (Amundsen, 1964), which is an equivalent correcting adjustment.[1]

Results for *e* based on the FFF model have generally been disregarded in the literature, mainly on the grounds of the strong condition of additive separability (preference independence) that is required for the approach to be valid. Quite clearly, Stern (1977) and Deaton and Muellbauer (1980) regarded this condition as unreasonably stringent. In the absence of additive separability, *e* has no meaning: for each monotonic transformation of the underlying preference function, a different value of *e* would emerge. However, in the case of food (and some other broadly defined product groups), Fellner (1967), Selvanathan and Selvanathan (1993) and Evans and Sezer (2002) have all argued that preference independence is a plausible assumption. Furthermore, Selvanathan (1988) tested the want-independence assumption for broad aggregates, using OECD data, and found it to be empirically valid. So the approach does seem worthy of some attention, despite the fact that both Cowell and Gardiner (1999) and Pearce and Ulph (1999) largely disregard empirical estimates of *e* based on the model. In fact, Pearce and Ulph argue that estimates of *e* based on the Fellner model are

[1]The Frisch formula can be expressed as follows:

$$e = \frac{(1 - wy)y}{p}$$

where y is the income elasticity of demand, p is the absolute value of the compensated price elasticity of demand and w is the budget share of the preference-independent product group.

The Amundsen modification is as follows:

$$e = \frac{by}{p}$$

where: b is the marginal propensity to spend on consumer goods other than the product group in question, so $b = (1–b')$, where b' is the marginal propensity to spend on the selected preference-independent product group; y is the income elasticity of demand; and p is the absolute value of the compensated price elasticity of demand.

It can easily be shown that $b = 1–wy$, meaning that the two approaches are equivalent.

overstated but they do not go on to state that this upward bias is easily corrected by applying the Frisch elasticities formula (Frisch, 1959).

The algebraic form of the demand model should, at the level of the individual consumer, reflect an underlying preference function in which the product in question enters in additively separable fashion. However, such a derivation of the demand equation would not necessarily survive aggregation to market level. So, experimentation with market demand models is required, with the preferred model being the one that best fits the data subject to the plausibility and statistical significance of the estimated income and price elasticities. Cointegration techniques, given large data samples, now make it easier to select appropriate market demand equations for the purpose of estimating statistically valid elasticities and associated *e* values. For example, Kula estimated a plausible *e* value for India using a constant elasticities model (CEM) that yielded a cointegrating relationship between the variables (see Kula (2002 and 2004)). Likewise, Evans and Sezer (2002) obtained virtually the same *e* value of 1.6 for the UK using the CEM. Indeed, recent work by Evans (2004a) relating to the UK and covering a longer data period (1965–2001) yielded the same estimate of 1.6 for *e*, with the CEM outperforming the AIDS model ('almost ideal demand system'). Further support for the relevance of constant elasticities comes from a pooled cross-section and time-series demand for food model, estimated in connection with the National Food Survey.[2] For 'all foods', it was found that the income elasticity of demand remained broadly constant over the period 1979–2000 (see MAFF (2000)).

Despite the good results for both the UK and India, there must still be some concern over estimates of *e* derived from the FFF model. Historically, *e* estimates for countries have been sensitive to the form of the demand model used (see, for example, Evans and Sezer (2002), especially table 2, page 1930). Also, recently reported work on France by Evans (2004b) suggests that the AIDS model outperforms the CEM, with tests showing that a cointegrating relationship between variables is only supported in the former case. An *e* estimate of 1.3 was obtained from the AIDS model, as opposed to the higher estimate of 1.8 from the CEM. Now, this latter estimate of *e* is similar to the CEM estimate of *e* for the UK, despite the fact that tests show it to be invalid; it is also very similar to the CEM estimate for India. Rather interestingly, the AIDS model results for the UK yield an *e* value of 1.25, a result that is very close to the French estimate of 1.3 (see Evans (2004a, table 2, page 445)). There is evidence here that the algebraic

[2]This evidence from the National Food Survey concerning the constancy of the income elasticity of demand for food is based on very large samples of household data. These cross-section data have been taken over three-year periods on a rolling basis from 1979 to 2000 (1979–81, 1980–82, 1981–83, etc. up to 1998–2000). See MAFF (2000, pages 97–98). In 2001, the National Food Survey was combined with the Family Expenditure Survey and renamed the Expenditure and Food Survey.

form of the demand model is having an important influence on the size of the derived *e* value. Furthermore, cointegrating relationships apply to different models in different countries. In part, no doubt, such inconsistency regarding model adequacy reflects the 'murky waters' of cointegration, given only modest-sized samples of time-series data.[3]

So, a mixed picture has emerged for *e* based on results from the application of the FFF model. The model is worthy of some cautious respect, but alternative methods of estimating *e* must be considered for the important purpose of cross-checking results.

3. Revealed social values

A suitable value for *e* may be revealed through government spending or tax policies. For example, the extent of progressiveness in a country's personal income tax rates can be viewed as a reflection of the government's degree of aversion to income inequality (a measure of *e*).[4] Stern (1977) and Cowell and Gardiner (1999) have produced estimates of *e* for the UK using personal income tax data. In both cases, the tax structure is assumed to be based on the principle of 'equal absolute sacrifice of satisfaction' and, in common with most researchers, iso-elastic utility functions are assumed. The model is set out formally below.

Assuming that personal income tax structures reflect, to an important extent, the principle of equal absolute sacrifice of satisfaction, then

$$U(Y) - U\left(Y - T(Y)\right) = K \,. \tag{3}$$

[3]Annual data were used to estimate the different demand models. While the use of quarterly data would increase the number of sample observations fourfold, it would not increase the effective sample size to anything like the same extent. This is especially the case in low-inflation periods, when successive quarters may yield very little variation in price data. Furthermore, a more complex dynamic model would become relevant in the context of quarterly data, with additional regressors including lagged variables. This increase in the number of variables may well introduce multicollinearity problems. Finally, the seasonal variation in the quarterly data is unlikely to exhibit regularity over a suitably lengthy data period and there is therefore a danger that the estimated regression model would not properly capture the seasonal effects. For all of these reasons, an annual model is preferred, even though its use necessitates only modest-sized data samples.

[4]Tax aversion in the case of individual taxpayers may lead to tax evasion problems. Underpayment of income tax is quite a serious problem in some of the OECD countries included in the data sample – for example, Italy, Spain and Turkey. This creates a divergence between a government's desired degree of progressiveness in the income tax structure, as expressed in the tax schedule, and the degree of progressiveness reflected in the actual tax payments. Clearly, more effective policing of the tax system is the answer, although this may prove costly. If serious tax evasion is a consequence of a lack of political will to remedy the situation, then there is a problem since the government is accepting the fact of a less progressive tax system than the one it has designed. In short, use of the official tax schedules would then produce an upward-biased measure of *e*. There is scope here for research into the extent of tax evasion in different countries and its persistence over time.

Furthermore, if utility functions are typically iso-elastic, then

(4) $$U(Y) = \frac{Y^{1-e} - 1}{1-e}.$$

In these two equations, Y is taxable income and $T(Y)$ is the income tax function reflecting the tax liabilities of an individual. Substituting equation (4) into equation (3) gives

(5) $$\frac{Y^{1-e} - 1}{1-e} - \frac{\left(Y - T(Y)\right)^{1-e} - 1}{1-e} = K.$$

Taking the total differential of equation (5) gives

(6) $$Y^{-e} - \left(Y - T(Y)\right)^{-e}(1-t) = 0$$

where t is the marginal tax rate. Simplifying equation (6) and taking logs gives

(7) $$\ln(1-t) = e\ln\left(1 - \frac{T(Y)}{Y}\right).$$

From equation (7), the following expression for e, a ratio of variables involving marginal and average tax rates, is obtained:

(8) $$e = \frac{\ln(1-t)}{\ln\left(1 - \frac{T(Y)}{Y}\right)}$$

where Y is taxable personal income, T is total income tax liability, T/Y is the average rate of income tax and t is the marginal rate of income tax.

Stern (1977) calculates the average tax rate as conventionally defined – that is, the ratio of tax liability to pre-tax income before the deduction of standard personal tax allowances. However, the use of this particular definition in order to calculate e from equation (8) imparts a strong upward bias to the value of e at relatively low levels of income. It is a better idea to calculate the average tax rate using pre-tax income after the deduction of standard tax allowances (although at low incomes where only a single tax rate applies, e would necessarily take a unitary value). This is the procedure adopted when estimating e values for 20 OECD countries in the next section of the paper. The economic justification for this calculation of the modified average tax rate is that people do not start paying tax on income until they have at least reached a subsistence wage level. It is only for income in excess of the subsistence level that it makes sense to argue that diminishing marginal utility applies.

From the tax-based results for *e* reported by Cowell and Gardiner (1999), it is not entirely clear whether average tax rates have been calculated in relation to income before, or after, the deduction of standard personal tax allowances. Their results for the UK, shown in Table 1, are based on regression analysis and relate to tax rates in 1999–2000. For the 'income tax only' model, they obtain a well-determined *e* estimate of 1.41. When employees' National Insurance contributions (NICs) are additionally included in the tax rate measures, then, as expected, a lower value of *e* is obtained (1.28). In both cases, the results for *e* are well above the UK Treasury's preferred value of unity. In fact, the estimate obtained from the 'income tax only' model is close to the Treasury's former preferred *e* value of 1.5 (see HM Treasury (1997)) and is similar to the *e* estimate obtained by Evans and Sezer (2002) using the adapted Fellner model.

The tax-based regression results from the 'income tax only' model seem more in keeping with the underlying theory concerning equal absolute sacrifice of satisfaction. From a government's viewpoint, the rationale for deductions relating to social security contributions is completely different from 'thoughts' concerning the optimum structure of income tax rates. The former are used, for example, to fund healthcare, and this is a 'needs-' rather than an income-related benefit. As such, they should not influence the calculation of *e*, meaning that the lower value of 1.28 obtained by Cowell and Gardiner (1999) should be regarded as a downward-biased estimate. Of course, from the individual's viewpoint, it can be argued that both are deductions from income. For the new work relating to 20 OECD countries, only income tax rates relating to wage earnings are considered.

Both Stern (1977) and Cowell and Gardiner (1999) conducted regressions based on equation (7) to estimate *e* but only after suppressing the intercept. For model-testing purposes, it would have been better not to have imposed the underlying theory in this way, so that a small and statistically insignificant constant would, for example, then offer empirical support for the model. Furthermore, a question might be raised concerning causation in the behavioural sense and thus the appropriateness of regression, especially as the effective sample size is small. Reversing the variables in equation (7) and then estimating the reciprocal of *e* might be more appropriate, since it makes more sense to argue that changes in marginal tax rates cause changes in average tax rates, rather than vice versa.

Recent work on tax-based estimates of *e* by Evans and Sezer (2004) involves direct calculation, using equation (8), at different discrete points in selected countries' gross wage earnings distributions for 2001. Six major OECD countries were considered, including the USA, the UK and Germany.

The estimates of *e* were concentrated in a reasonably narrow band: the highest value was 1.6 (Germany) and the lowest was 1.3 (France). Rather interestingly, the estimate of *e* for the UK was 1.5 (see Table 1). In Section III, many more countries are considered for the year 2002 and cross-country regressions are conducted. Greater care is taken with both the relevant tax data and the statistical testing of results. The simple 'equal absolute sacrifice' model appears to yield sensible estimates of *e*, although they are somewhat higher than the UK Treasury's preferred unitary value. The model is certainly worthy of more attention for the purpose of estimating *e* values for countries, especially as Stern (1977) found that empirically it outperforms the more complex models of income tax structures. In fact, the model's simplicity is its virtue, especially if statistical testing supports the underlying assumptions.

4. Concluding comments on the evidence

Estimates of *e* based on models of lifetime consumption behaviour have been taken most seriously at policy level in recent years. In particular, the decision of HM Treasury to select a unitary value for *e* was heavily influenced by the findings of Blundell, Browning and Meghir (1994). While both the theory and the econometrics in this particular study are impressive, it still seems unwise to place too much emphasis on results from a sample data period that finished nearly 20 years ago. Furthermore, the data period in question was certainly turbulent in terms of market shocks and major macroeconomic policy changes. An updated study for the UK (and other countries) is required for a more 'reliable' estimate of *e*.

Evidence on *e* based on an adapted version of the Fellner demand model for a preference-independent good (Fellner, 1967) has, perhaps, been unfairly disregarded in the literature. This is largely due to the strong assumption of additive separability in the utility function that the derivation of *e* requires. If this restriction is invalid, then a measure of *e* cannot be sensibly derived from the estimated income and price elasticities: the value of *e* would change with each monotonic transformation of the utility function. However, for broad product groups (especially food), the assumption of preference independence was tested by Selvanathan (1988) and found to hold. Recent empirical work based on this method has produced sensible and well-determined values of *e* in the context of cointegrating demand relationships (see Evans and Sezer (2002), Kula (2002 and 2004) and Evans (2004a)). It is not good news all round, however. There is some evidence to suggest that the value of *e* is sensitive to demand model specification and that different models for different countries yield cointegrating relationships (see Evans (2004a and 2004b)).

Apart from Stern (1977) and Cowell and Gardiner (1999), the revealed social values approach based on personal income tax structures has been under-utilised as a method of estimating *e*. The latter estimated an *e* value of 1.41 for the UK in 1999–2000 using data on income tax rates alone. This result is well above the Treasury's preferred value of unity, very close to its previous value of 1.5 (see HM Treasury (1997)) and not dissimilar to the estimate of 1.6 for the UK obtained by Evans and Sezer (2002) and Evans (2004a). A recent preliminary study involving a tax-based estimation of *e* revealed that for six major countries, estimates are in the range 1.3 to 1.6, with an *e* value of 1.5 for the UK (see Evans and Sezer (2004)). An extended and more rigorous tax-based cross-country analysis involving 20 OECD countries is reported in the next section of the paper.

III. New research findings on *e*

The equal absolute sacrifice model, assuming an iso-elastic social utility function, is tested for 20 OECD countries using income tax rates on the gross wage earnings of single persons without dependants in 2002. The required wage and tax data are supplied by OECD (2003). In the first instance, the value of *e* is calculated directly for each country at different wage levels, and then tests are conducted to see if constancy of *e* is, on average, supported statistically. Then cross-country regressions are conducted in order to estimate values of *e* using pooled income data based on these same contrasting wage levels. Tests employing dummy variables are carried out to determine whether or not there is statistical support for iso-elastic preferences. The direct calculation and regression approaches to estimating *e* are based on equations (8) and (7) respectively. After a proper consideration of the data-set used, the results are reported, analysed and discussed.

1. Selection of countries and tax rate data

Some OECD countries are excluded from the sample. For example, countries where most of the revenue is generated from state and local taxes rather than central government taxes are omitted; these include Denmark, Sweden and Finland. Provided at least 50 per cent of tax revenue with respect to wages is raised from central government taxes, then it is not unreasonable to consider including countries in the sample. In the case of the USA, only federal taxes are considered owing to the inter-state variation in state and local tax rates.

Out of the 20 countries included in the sample, a few had tax structures giving rise to relatively large differences between *e* values at contrasting wage levels – for example, Ireland, the UK and Australia. The exclusion of these countries from the sample would produce very strong support for iso-elastic preferences. However, such data-mining practices must be rejected in the interests of legitimate statistical testing procedures. See Table 2 for evidence of differences in *e* values by income level for all 20 OECD countries.

TABLE 2

Tax-based calculations of e *for 20 OECD countries at 'high' and 'low' income levels (2002)*

	t	T/Y	$-ln(1-t)$	$-ln(1-T/Y)$	e
Australia 45,851					
Low 35,000	0.30	0.210	0.35667	0.23572	1.51
High 55,000	0.42	0.258	0.54473	0.29841	1.82
Austria 18,345					
Low 14,530	0.31	0.187	0.37106	0.20702	1.79
High 29,070	0.41	0.269	0.52763	0.31334	1.68
Belgium 24,979					
Low 20,985	0.45	0.361	0.59784	0.44785	1.33
High 36,560	0.50	0.409	0.69315	0.52594	1.32
Canada 38,568					
Low 47,515	0.22	0.180	0.24846	0.19845	1.25
High 83,177	0.26	0.207	0.30110	0.23193	1.30
Czech Rep. 142,247					
Low 163,800	0.20	0.167	0.22314	0.18272	1.22
High 274,800	0.25	0.191	0.28768	0.21196	1.36
France 18,467					
Low 17,043	0.21	0.151	0.23572	0.16370	1.44
High 28,680	0.31	0.248	0.37106	0.28502	1.30
Germany 30,145					
Low 29,049	0.29	0.223	0.34249	0.25231	1.36
High 100,000	0.485	0.386	0.66359	0.48776	1.36
Hungary 1,056,835					
Low 900,000	0.30	0.233	0.35667	0.26527	1.34
High 2,000,000	0.40	0.310	0.51083	0.37106	1.38
Ireland 25,330					
Low 14,000	0.20	0.200	0.22314	0.22314	1.00
High 56,000	0.42	0.310	0.54473	0.37106	1.47
Italy 19,493					
Low 23,420	0.32	0.241	0.38566	0.27575	1.40
High 50,355	0.39	0.310	0.49430	0.37106	1.33
Japan 2,057,989					
Low 2,650,000	0.18	0.140	0.19845	0.15082	1.32
High 5,150,000	0.28	0.187	0.32850	0.20702	1.59

Table 2 continues overleaf

TABLE 2 continued

	t	T/Y	–ln(1–t)	–ln(1–T/Y)	e
New Zealand 39,411					
Low 49,000	0.33	0.225	0.40048	0.25489	1.57
High 90,000	0.39	0.293	0.49430	0.34672	1.43
Norway 218,800					
Low 575,000	0.415	0.340	0.53614	0.41551	1.29
High 1,000,000	0.475	0.382	0.64436	0.48127	1.34
Poland 19,201					
Low 55,536	0.30	0.227	0.35667	0.25748	1.38
High 92,560	0.40	0.276	0.51083	0.32296	1.58
Portugal 5,318					
Low 10,800	0.24	0.175	0.27444	0.19237	1.43
High 25,000	0.34	0.250	0.41551	0.28768	1.44
Slovakia 95,263					
Low 135,000	0.20	0.133	0.22314	0.14272	1.56
High 288,000	0.28	0.199	0.32850	0.22189	1.48
Spain 9,630					
Low 8,276	0.24	0.213	0.27444	0.23953	1.15
High 19,004	0.283	0.242	0.33268	0.27707	1.20
Turkey 8,087m					
Low 6,650m	0.25	0.203	0.28768	0.22690	1.27
High 14,250m	0.30	0.251	0.35667	0.28902	1.23
UK 15,093					
Low 14,950	0.22	0.205	0.24846	0.22941	1.08
High 59,900	0.40	0.306	0.51083	0.36528	1.40
USA[a] 24,488					
Low 16,975	0.15	0.132	0.16252	0.14156	1.15
High 47,825	0.27	0.194	0.31471	0.21567	1.45

[a]Only central government tax rates have been considered for the USA because of the inter-state variation in state and local tax rates.

Notes:

1. Pre-tax annual gross earnings figures for 2002, less standard tax allowances, are expressed in units of national currency for each of the countries. 'Low' and 'high' income figures are shown, along with taxable wage earnings at the average production wage (APW) in manufacturing industries.
2. For each band of taxable gross earnings subject to a specific marginal tax rate, t, the adjusted average tax rate, T/Y, has been computed at the midpoint of the band. Y is defined as pre-tax gross annual earnings (2002) less standard tax allowances.

Source: OECD, 2003.

The selection of the contrasting wage levels at which marginal and average tax rates are calculated, for the purpose of estimating e, depends on the number of different tax rates and the breadth of the taxable income bands. For countries with 'lumpy' tax structures and broad income bands in which a particular marginal tax rate applies, the difference between selected 'low' and 'high' levels of taxable income is correspondingly large; see, for example, Ireland and the UK in Table 2. As far as possible, the 'low' income level is selected so that it is close to a country's average production wage (APW). The 'high' income level is then selected to be substantially above the APW. Table 2 gives full details of the relevant levels of taxable income for each country and the APW, less standard tax allowances, for reference purposes. The smallest factor by which the 'high' income level exceeds the 'low' income level is 57 per cent (Australia). Correspondingly, the largest margin is 300 per cent (Ireland). Such differences in margins are unavoidable, given the small number of different marginal tax rates in some countries and the need to calculate average tax rates at the midpoints of the relevant taxable income bands.

The calculation of average tax rates at the midpoints of taxable income bands largely avoids the problem of biased calculations or estimates of *e*. For example, if the average tax rate were calculated at set percentages of the APW for all countries, then in some cases it might correspond to a wage level just below the threshold at which the marginal tax rate rises. In other cases, the wage level might be just above the relevant threshold. In the former case, a downward-biased estimate of *e* would be obtained, while an upward-biased estimate would be the result in the latter case. Mostly, calculating average tax rates at the midpoints of taxable income bands is a procedure that should avoid any serious bias in the estimation of *e*. However, where income bands are unavoidably wide, including in a few cases open-ended groups, there is greater scope for error. The only way to avoid such error completely is to calculate the actual average tax rate for all single persons (without dependants) paying the same marginal rate of income tax. In order to do this, knowledge of total taxable wage income and total tax liabilities, with respect to wage earnings, would be required for the relevant marginal tax rate bands in all 20 countries. In short, this would necessitate obtaining considerably more detailed tax data from the tax authorities for each country. The OECD is unable to supply this level of information, so for current purposes midpoint approximations of the relevant average tax rates are taken. In most cases, this procedure should be reasonable enough.

All the relevant data on marginal and average tax rates on wages are detailed for each country in Table 2. The relevant log values of the tax rate variables, consistent with equation (8), are also shown, along with directly calculated *e* values at both 'low' and 'high' wage levels.

2. Analysis of results for *e*

The tests relating to the cross-country *e* values are concerned to measure mean values at both 'low' and 'high' income levels and to determine whether or not there is a statistically significant difference between these means. Testing for constancy of *e* is especially important because the model

TABLE 3

Tax-based e *values for 20 OECD countries at high and low income values*

	e_1 (*low* Y)	e_2 (*high* Y)	e_2–e_1	$(e_2–e_1)^2$
Australia	1.51	1.82	0.31	0.0961
Austria	1.79	1.68	–0.11	0.0121
Belgium	1.33	1.32	–0.01	0.0001
Canada	1.25	1.30	0.05	0.0025
Czech Republic	1.22	1.36	0.14	0.0196
France	1.44	1.30	–0.14	0.0196
Germany	1.36	1.36	0.00	0.0000
Hungary	1.34	1.38	0.04	0.0016
Ireland	1.00	1.47	0.47	0.2209
Italy	1.40	1.33	–0.07	0.0049
Japan	1.32	1.59	0.27	0.0729
New Zealand	1.57	1.43	–0.14	0.0196
Norway	1.29	1.34	0.05	0.0025
Poland	1.38	1.58	0.20	0.0400
Portugal	1.43	1.44	0.01	0.0001
Slovakia	1.56	1.48	–0.08	0.0064
Spain	1.15	1.20	0.05	0.0025
Turkey	1.27	1.23	–0.04	0.0016
UK	1.08	1.40	0.32	0.1024
USA	1.15	1.45	0.30	0.0900
Sums			1.62	0.7154

Note: The *e* values shown in this table are those calculated from the tax rate data in Table 2.

generating equation (8) assumes iso-elastic social preferences. The tests are conducted on unweighted data for both *e* and the relevant tax rate variables (see Tables 2 and 3). This is a reasonable procedure in view of the fact that for most of the 20 countries, *e* values calculated at both low and high income levels are clustered in a narrow band (see Table 3). Furthermore, in the case of the outlier values, there is no tendency for the relatively small or large values to be associated with the size or economic importance of a country. For example, Ireland and the UK both have relatively low *e* values at 'low' income levels.

The initial tests, reported in part (a) of this section, are matched samples tests based on the calculated *e* values highlighted in Table 3. These 'low-' and 'high-'income *e* values have been calculated directly for each country using the tax rate information detailed in Table 2. The calculations are based on the direct application of equation (8). It will be shown that both parametric and non-parametric tests offer support for constancy of *e*.

In part (b), regression analysis is used to estimate cross-country *e* values. The relevant tax rate data, shown in Table 2, are fitted to equation (7). The data for both wage levels are pooled, so that dummy variables can be employed to test the validity of the constancy of *e* restriction. The 'best' estimate(s) of *e* is/are obtained from the most appropriate model specification, as determined by the data. It will be shown that there is statistical support, on average, for constancy of *e* and that plausible and well-determined *e* values are yielded by the regression analysis.

(a) Matched samples tests

A parametric test based on the t-distribution and the non-parametric Wilcoxon signed ranks test are both used to test for constancy of *e*, on average, at different wage levels. The former test has more power, but since the sample of countries is not a random selection, it is not entirely appropriate.[5] The non-parametric test may lack power (although it is more powerful than the matched samples sign test) but is more appropriate. If both tests offer statistical support for constancy of *e* then this can be taken as sufficiently strong evidence.

Table 3 shows the differences between low- and high-wage *e* values for 20 OECD countries. The mean low-wage *e* value for these countries is 1.34 and the mean high-wage value is 1.42 (6 per cent higher). Is the mean difference between the high- and low-wage *e* values significantly different from zero? Full details of the matched samples t-test are shown in Appendix A and the results reveal that since calculated t is marginally smaller than the critical t-value at the 5 per cent significance level, then the hypothesis of iso-elastic social preferences can just be accepted at the 5 per cent level.

Details of the results for the Wilcoxon signed ranks test are also shown in Appendix A. This test considers the frequency of both the positive and negative differences in *e* values, across countries, and ranks the differences according to their absolute size. The sum of the ranks for the positive (or negative) differences is then compared with the mean of the sums of the ranks for the positive and negative differences in *e* values. Full details are shown in Appendix A. The calculated Z-score of –1.53 means that, on average, constancy of *e* cannot be rejected at even the 10 per cent level.

So, both the parametric and non-parametric tests employed offer statistical support for constancy of *e*. Furthermore, the average *e* value taken for all 20 countries, over both low and high wage levels, is close to 1.4 and this does seem a plausible result.

[5]Differences in *e* values are, by appeal to the central limit theorem, likely to be normally distributed. So, if the sample had consisted of randomly selected observations then the matched samples t-test would have been appropriate. Although the data for 20 OECD countries do not constitute a random sample, the power of the test is still appealing. In any case, other tests are also applied to the data.

(b) Regression analysis

Cross-country regression estimates of *e* based on pooled 'low' and 'high' income data are shown in Table 4. This procedure permits the testing of the assumption concerning constancy of *e*. Specification 1 is the restricted model imposing constancy, and specification 3 is the unrestricted model including shift and slope dummy variables. Specifications 4 and 5 are partially restricted models. The regressions are based on equation (7) but initially an intercept is included so that the underlying theory is not merely imposed from the outset. In both specifications 1 and 3, the estimated intercepts are relatively small and statistically insignificant.

Although both Stern (1977) and Cowell and Gardiner (1999) estimated their tax-based regressions in accordance with Table 4, there is a case for reversing the implied causality and estimating the reciprocal of *e*. So, changes in tax rate variables involving marginal rates are assumed to cause changes in tax variables involving average tax rates. This line of thinking on causal link seems more plausible and, as such, Table 5 reverses all the regressions shown in Table 4, thus providing estimates of 1/*e*. Once again, the estimated intercepts in specifications 1 and 3 are both relatively small and statistically insignificant. The regression coefficients are well determined and the explanatory power of the equations is, in each case, close to 90 per cent.

For both the direct and indirect estimates of *e*, presented in Tables 4 and 5 respectively, an F-test reveals that the unrestricted model (specification 3 in both tables) fails to explain significantly more variation in the dependent

TABLE 4

Pooled tax-based regressions: testing constancy of e *for 20 OECD countries (2002)*

Dependent variable:	*1.* *ln(1–*t)	*2.* *ln(1–*t)	*3.* *ln(1–*t)	*4.* *ln(1–*t)	*5.* *ln(1–*t)
Constant	0.0066		0.0073	0.0120	0.0238
	(0.30)		(0.24)	(0.57)	(1.02)
ln(1–*T*/*Y*)	1.36	1.38	1.31	1.29	1.25
	(18.57)	(59.23)	(10.41)	(16.08)	(12.73)
Dummy			0.038	0.029	
			(0.81)	(1.90)	
*D*ln(1–*T*/*Y*)			–0.034		0.091
			(–0.21)		(1.71)
RSS	0.07385	0.07400	0.06722	0.06730	0.06844
Adjusted R^2	0.898		0.902	0.905	0.903
RESET (F)	1.95		0.95		
	[4.10]		[4.13]		

Notes: Dummy = 1 for 'high' income observations; otherwise 0. *D* = 1 for 'high' income observations; otherwise 0. Figures in parentheses are t-ratios. RESET = Ramsey's regression specification test (F-version). Critical 5 per cent F-values are shown in square brackets.

TABLE 5

*Pooled tax-based regressions: testing constancy of 1/*e *for 20 OECD countries (2002)*

Dependent variable:	*1.* *ln(1–*T/Y*)*	*2.* *ln(1–*T/Y*)*	*3.* *ln(1–*T/Y*)*	*4.* *ln(1–*T/Y*)*	*5.* *ln(1–*T/Y*)*
Constant	0.0234		0.0250	0.0213	0.0186
	(1.60)		(1.15)	(1.42)	(1.11)
ln(1–*t*)	0.662	0.716	0.667	0.679	0.685
	(18.57)	(59.23)	(10.24)	(16.08)	(13.21)
Dummy			–0.016	–0.008	
			(–0.47)	(–0.74)	
*D*ln(1–*t*)			0.020		–0.018
			(0.24)		(–0.61)
RSS	0.03595	0.03840	0.03537	0.03543	0.03559
Adjusted R^2	0.898		0.894	0.897	0.896
RESET (F)	2.28		1.36		
	[4.10]		[4.13]		

Notes: See Notes to Table 4.

variable than the restricted model (specification 1 in both cases). For details of the F-test results, see Appendix B. So, the joint insignificance of the shift and slope dummy variables in specification 3 in both tables offers statistical support for the validity of the restricted model and thus the constancy of *e*. As a precaution, the restricted model is also tested against partially restricted models (specifications 4 and 5 in both tables). Inspection of the t-ratios reveals that neither the shift nor slope dummy variable is statistically significant at the 5 per cent level. However, in Table 5, the relevant dummy variable coefficients are highly insignificant and negligible in value. This finding is, taken by itself, a reason for preferring estimates of *e* yielded by the indirect model.

Tests for heteroscedasticity in relation to the restricted models allow the acceptance of homoscedastic disturbances at the 5 per cent level, although support for homoscedasticity is stronger in the case of the direct model (specification 1 in Table 4). (The Glejser and Spearman rank correlation tests were used. Details of the results can be provided on request to the author.) Ramsey's RESET test for functional form misspecification clearly supports the equal absolute sacrifice model with respect to the OECD data sample. The F-test results for both the unrestricted and restricted models (specifications 3 and 1 respectively) reveal that adding the squares of the fitted values of the dependent variable as an additional explanatory variable fails, in each case, to increase significantly the explanatory power of the regression equations. For both models, the calculated F-values are clearly below the critical 5 per cent F-values.

While the restricted models receive empirical support, with both accounting for 90 per cent of the variation in the dependent variable, they still produce differing but well-determined estimates of *e*: 1.36 in the case of the direct model and 1.51 from the indirect model. This difference appears to be caused by the larger but still statistically insignificant constant term in the indirect model (compare the results for specification 1 in Tables 4 and 5). If the constant term is dropped from the restricted form of both models, then the resulting direct and indirect estimates of *e* are both close to 1.4 (see specification 2). This value is taken as the preferred result.

IV. Policy implications

First, the tax-based estimates of *e* reported in this paper are used to help estimate appropriate social discount rates for five major countries; these are discount rates measured on a consistent time preference basis. As shown in Table 6, these rates are quite similar in value, ranging from 3.7 per cent to 4.4 per cent. This is followed by a consideration of the role of distributional welfare weights in the determination of regional policy funding decisions in a future Europe characterised by harmonised fiscal and monetary policies. The sizes of these weights are significantly influenced by the values of *e*.

1. Measures of *e* and social discount rates

The tax-based *e* values from this study can be used for the purpose of calculating social time preference rates for selected countries. Inspection of Table 3 reveals that for five major OECD countries (France, Germany, Japan, the UK and the USA), their average ‘low-’ and ‘high-’wage *e* values range between 1.25 (UK) and 1.45 (Japan). The midpoint of this reasonably tight band, 1.35, is very close to the preferred *e* value of 1.4 for the full sample of 20 OECD countries. As such, it will be assumed for current purposes that an *e* value of 1.35 applies to all five countries.

From the ‘Ramsey’ equation for the stpr (see equation (1)), suitable figures are also required for both *p* and *g*. For the latter, average annual per-capita growth figures for real consumer spending, over the period 1970–2002, are taken from the OECD National Accounts (see OECD (2004)). Values of *p* are based on recent average death rates for the five countries. For justification of a mortality-based utility discount rate, see, for example, Kula (1987) and Pearce and Ulph (1999). In each case, the death rates for 2002 and 2003 are very close to 1 per cent, so for all five countries this particular value serves as an appropriate measure of *p*.

The full set of results for the five countries, including the final stpr figures, are shown in Table 6. Given the reasonable assumptions made, it is

TABLE 6

Social time preference rates for five major countries

	p (%)	g (%)	e	stpr (%)
France	1.0	2.0	1.35	3.7
Germany	1.0	2.2	1.35	4.0
Japan	1.0	2.5	1.35	4.4
UK	1.0	2.1	1.35	3.8
USA	1.0	2.2	1.35	4.0

Notes: Mortality-based measures are used for *p* (approximate average annual death rate in recent years). Measures of *g* are based on average annual real per-capita consumption growth over the period 1970–2002. The *e* measure of 1.35 is based on the tax-based estimates presented in this paper.

only the per-capita growth rates, *g*, that differ between countries and the variation involved is expectedly small. The stpr figures range between 3.7 per cent for France and 4.4 per cent for Japan. In the case of France, the stpr is less than half of the 8 per cent rate that has usually been applied in the appraisal of social projects (see, for example, OECD (2001)). So, the application of the stpr in France should result in a more generous allocation of government funds to long-term social projects yielding sizeable net social benefits to future generations. Furthermore, this consistent approach to estimating stprs for the five major countries would result in the application of very similar social discount rates. This is unlike the current situation with official rates based on rather different methods.

2. Welfare weights and the funding of regional policy options in a future Europe

The tax-based estimates of *e* could be used to help construct a series of distributional welfare weights for application in EU regional policy. Confidence in the appropriateness of these estimates would be further reinforced if updated work on lifetime consumption behaviour were to yield a similar set of results for *e*, in the range 1.3 to 1.5. The construction of such welfare weights for application in regional social project and policy appraisals in Europe would certainly be a viable policy option in a more integrated Europe in which both monetary and fiscal harmony are fully realised. Although such accord still seems a rather distant prospect, the future possibility of applying appropriate regional welfare weights based on relative per-capita real consumption levels in countries (see equation (2)) is still very real. Not only would it be possible to construct relative consumption (or income) levels consistently in euros, but also harmonised tax rates would, given iso-elastic preferences, be consistent with a single communal *e* value. These two factors of consistency make it possible to derive an appropriate set of distributional welfare weights. Such weights could, for example, be applied in regional social project appraisals and thus significantly influence the allocation of official funds across countries. Moreover, these regional welfare weights could be used as the basis for determining the net contributions that each member country should make to the EU budget. An equitable set of welfare weights, applied appropriately, may well serve to help reduce regional imbalances within the EU. Such a reduction could, in turn, help to alleviate any long-term migration problems and help promote economic growth in Europe.

V. Concluding comments

The tax-based estimates of *e* produce a best average estimate for countries that is close to 1.4. On the strength of the pooled income regression results for the preferred restricted model (see specification 1 in Table 4), the lower and upper 95 per cent confidence limits for average *e* are 1.21 and 1.51 respectively. Significantly, alternative methods of estimating *e* seem to produce values for countries that are either within this range or just outside. See, for example, Evans and Sezer (2002) and Kula (2004), who both employed an adapted Fellner model to estimate *e* values of approximately 1.6 for the UK and India respectively. In fact, Evans (2004b) used this same model to estimate an *e* value of 1.3 for France. Even the standard Blundell et al. model of lifetime consumption behaviour produced an average *e* value of 1.3, an estimate that is close to the tax-based *e* values estimated in this paper (see Blundell, Browning and Meghir (1994)).

The recent empirical evidence on *e* is much more in keeping with the UK Treasury's preferred former value of 1.5 rather than its current preference of 1.0 (see HM Treasury (1997 and 2003)). In fact, previous support for the value of 1.5 in the case of the UK was based on saving behaviour; see, for example, Scott (1989). However, while evidence against an *e* value as low as unity is certainly mounting, there is a definite need for further work based on lifetime consumption behaviour in order to help clinch matters. In the case of the UK, for example, evidence drawn from this latter approach urgently needs updating so that it reflects consumer behaviour in an era of low inflation and deregulated financial markets. If estimates of *e* based on both a tax-based approach and lifetime consumption behaviour turn out to be similar for countries, in the region of 1.3 to 1.5, then the UK Treasury would need to revise upwards its preferred value of *e*.

Appendix A. Matched samples tests based on information in Tables 2 and 3

1. T-test

Let $e_2 – e_1 = d$ (see Table 3). Then

H_0: $\mu_d = 0$ (accept iso-elastic social utility function);
H_1: $\mu_d \neq 0$ (reject iso-elastic social utility function).

Calculated $t = (d - 0) \big/ \sqrt{s^2 / (n-1)} = 0.081 \big/ \sqrt{0.029209/19} = 2.07$.
Critical t_{19} (5 per cent significance level) = 2.093 .
Thus H_0 is just accepted at the 5 per cent significance level.

2. Non-parametric Wilcoxon signed ranks test

The highest rank (20) is applied to the largest absolute value of $e_2 – e_1$. The mean of the sums of the ranks for positive and negative values of $e_2 – e_1$ is 105. So

$$\mu_R = \frac{n(n+1)}{4} = 105 .$$

The standard deviation of the ranks is

$$\sigma_R = \sqrt{\frac{n(n+1)(2n+1)}{24}} = 26.8 .$$

The following Z-value can be calculated:

$$Z = \frac{R' - \mu_R}{\sigma_R} = \frac{64 - 105}{26.8} = -1.53$$

where R' is the sum of the ranks for negative values of $e_2 – e_1$. Calculated Z is not a statistically significant result at either the 5 per cent or 10 per cent significance level. So, on average, for countries, the hypothesis of constancy of *e* can be accepted.

Appendix B. F-tests for constancy of *e*

These tests focus on specifications 1 and 3 in the pooled regressions presented in Tables 4 and 5, where Table 4 provides direct estimates of *e* and Table 5 provides estimates of the reciprocal of *e*. In each case, specification 3 is the unrestricted model including shift and slope dummies for the 'high' income observations, and specification 1 is the fully restricted model in which constancy of *e* is imposed. The following F-test is conducted to test the validity of the restriction:

$$F = \frac{SS_U - SS_R}{K_U - K_R} \bigg/ \frac{RSS_U}{n - K_U - 1}$$

where SS_U is the explained sum of squares in the unrestricted model, SS_R is the explained sum of squares in the restricted model, K_U is the number of explanatory variables in the unrestricted model, K_R is the number of explanatory variables in the restricted model, RSS_U is the residual sum of squares in the unrestricted model and *n* is the number of sample observations in the pooled model.

1. Test based on direct estimate of *e* (Table 4)

Calculated F = $\frac{0.00663}{2} \bigg/ \frac{0.06722}{36} = 1.78$.

Critical $F_{2,36}$ (5 per cent significance level) = 3.26 .

2. Test based on indirect estimate of *e* (Table 5)

Calculated F = $\frac{0.00058}{2} \bigg/ \frac{0.03537}{36} = 0.30$.

Critical $F_{2,36}$ (5 per cent significance level) = 3.26 .

In both cases, the restricted model is supported at the 5 per cent significance level and the null hypothesis of constancy of *e* can be comfortably accepted.

References

Amiel, Y., Creedy, J. and Hurn, D. (1999), 'Measuring attitudes towards inequality', *Scandinavian Journal of Economics*, vol. 101, pp. 83–96.

Amundsen, A. (1964), 'Private consumption in Norway 1930–1978', in J. Sandee (ed.), *Europe's Future Consumption*, Amsterdam: North Holland Publishing Company.

Attanasio, O. P. and Browning, M. (1995), 'Consumption over the life cycle and over the business cycle', *American Economic Review*, vol. 85, pp. 1118–37.

Barsky, R. B., Kimball, M. S., Juster, F. T. and Shapiro, M. D. (1995), 'Preference parameters and behavioural heterogeneity: an experimental approach in the Health and Retirement Survey', National Bureau for Economic Research, Working Paper no. 5213.

Besley, T. and Meghir, C. (1998), *Tax Based Savings Incentives*, Technical Report, London: London School of Economics.

Blue, E. N. and Tweeten, C. (1997), 'The estimation of marginal utility of income for application to agricultural policy analysis', *Agricultural Economics*, vol. 16, pp. 155–69.

Blundell, R., Browning, M. and Meghir, C. (1994), 'Consumer demand and the life-cycle allocation of household expenditures', *Review of Economic Studies*, vol. 61, pp. 57–80.

Cowell, F. A. and Gardiner, K. (1999), 'Welfare weights', London School of Economics, STICERD, Economics Research Paper no. 20.

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

Evans, D. (2004a), 'The elevated status of the elasticity of marginal utility of consumption', *Applied Economics Letters*, vol. 11, pp. 443–7.

— (2004b), 'A social discount rate for France', *Applied Economics Letters*, vol. 11, pp. 803–8.

—, Kula, E. and Sezer, H. (2005), 'Regional welfare weights for the UK: England, Scotland, Wales and Northern Ireland', *Regional Studies*, forthcoming.

— and Sezer, H. (2002), 'A time preference measure of the social discount rate for the UK', *Applied Economics*, vol. 34, pp. 1925–34.

— and — (2004), 'Social discount rates for six major countries', *Applied Economics Letters*, vol. 11, pp. 557–60.

Fellner, W. (1967), 'Operational utility: the theoretical background and a measurement', in W. Fellner (ed.), *Ten Economic Studies in the Tradition of Irving Fisher*, New York: John Wiley and Sons.

Fisher, I. (1927), 'A statistical method for measuring marginal utility', in *Economic Essays Contributed in Honour of J. Bates*, London: Macmillan.

Frisch, R. (1932), *New Methods of Measuring Marginal Utility*, Tubingen: J. C. B. Mohr.

— (1959), 'A complete system for computing all direct and cross demand elasticities in a model with many sectors', *Econometrica*, vol. 27, pp. 177–96.

HM Treasury (1997), *Appraisal and Evaluation in Central Government*, The Green Book, London: HMSO.

— (2003), *Appraisal and Evaluation in Central Government*, The Green Book, London: HMSO.

Hotelling, H. (1931), 'Economics of exhaustible resources', *Journal of Political Economy*, vol. 39, pp. 137–75.

Kula, E. (1987), 'Social interest rate for public sector appraisal in the United Kingdom, United States and Canada', *Project Appraisal*, vol. 2, pp. 169–74.

— (2002), 'Regional welfare weights in investment appraisal: the case of India', *Journal of Regional Analysis and Policy*, vol. 32, pp. 99–114.

— (2004), 'Estimation of a social rate of interest for India', *Journal of Agricultural Economics*, vol. 55, pp. 91–9.

MAFF (2000), *The National Food Survey*, London: Office for National Statistics.

OECD (2001), *Assessing the Benefits of Transport*, European Conference of Ministers of Transport.

— (2003), *Taxing Wages 2002*, Paris: Organisation for Economic Cooperation and Development.

— (2004), *Extracts from the National Accounts of OECD Countries Database*, Paris: Organisation for Economic Cooperation and Development.

OXERA (2002), *A Social Time Preference Rate for Use in Long-Term Discounting: A Report for ODPM, DfT and DEFRA*, Oxford.

Pearce, D. and Ulph, D. (1995), 'A social discount rate for the UK', University of East Anglia, School of Environmental Studies, CSERGE, Working Paper no. 95-01.

— and — (1999), 'A social discount rate for the UK', in D. W. Pearce (ed.), *Environmental Economics: Essays in Ecological Economics and Sustainable Development*, Cheltenham: Edward Elgar.

Ramsey, F. P. (1928), 'A mathematical theory of saving', *Economic Journal*, vol. 38, pp. 543–59.

Scott, M. F. G. (1989), *A New View of Economic Growth*, Oxford: Clarendon Press.

Selvanathan, S. (1988), 'A system-wide analysis of international consumption patterns: advanced econometric series', Ph.D. thesis, University of Western Australia.

— and Selvanathan, E. A. (1993), 'A cross-country analysis of consumption patterns', *Applied Economics*, vol. 25, pp. 1245–59.

Slade, M. E. (1982), 'Trends in natural-resource commodity prices: an analysis of time domain', *Journal of Environmental Economics and Management*, vol. 9, pp. 122–37.

Stern, H. N. (1977), 'Welfare weights and the elasticity of marginal utility of income', in M. Artis and R. Norbay (eds), *Proceedings of the Annual Conference of the Association of University Teachers of Economics*, Oxford: Blackwell.