1. Introduction

Studying the causes and correlates of human happiness has become one of the hot topics in economics over the last decade, with both the size and depth of the literature increasing at an exponential rate. As an illustration, a search of ECONLIT for journal articles that have either ‘Happiness’, ‘Life Satisfaction’ or ‘Well-being’ in the title, shows that there were 417 articles identified as having been published between 1960 and 2005. Of these 315 (79%) have been published since 1995, 237 (59%) have been published since 2000, and one-third of the literature (31%, or 125 articles) has appeared in print in just the last three years. Focusing only on the period 2000-2005, this measure of economists’ interest in the ‘economics’ or ‘science’ of happiness, would place that interest at about half the size of the interest in ‘Wages’ and a just a little below ‘Discrimination’. However, ‘happiness’ is quickly catching up; it is in its ‘industrial revolution’ stage. A search for Working Papers in SSRN or RePEc clearly confirms the booming interest in the economics of happiness.

The range of issues that have been addressed in this literature is large. However, two topics have attracted the most interest: (1) the relationship between income and happiness; and (2) the extent to which unemployment reduces happiness. While it is now widely accepted that unemployment reduces well-being, even after controlling for the associated movements in income, the relationship between income and happiness remains more contentious.

---

1 Clark: PSE, Paris, France; Frijters: School of Economics and Finance, Queensland University of Technology, Australia; Shields: Department of Economics, University of Melbourne, Australia. We are grateful to the editor and two referees for valuable comments.

2 Throughout this review we will use the terms well-being, happiness and life satisfaction interchangeably. However, when relevant we will highlight the actual measure of happiness used in the studies that we will review.

3 This only covers articles published until July 2005. Search of ECONLIT was conducted on the 14th November, 2005. Only journal articles written in English were included in the search. Of the 135 articles in total, 116 (87%) have ‘Happiness” in the title. There were only three articles with “Life Satisfaction” in the title before 1995.
One of the key catalysts in this literature were the studies by Richard Easterlin (1974 and 1995), underlining the ‘paradox’ of substantial real income growth in Western countries over the last fifty years, but without any corresponding rise in reported happiness levels. To borrow a term from health economics, it looks as if individuals in rich countries are ‘flat of the curve’, with additional income buying little if any gains in happiness. It has been strongly argued that once an individual rises above the poverty line, the main source of increased well-being is not income but rather friends and a good family life (see, for example, Robert Lane, 2001). The radical implication for developed countries is that economic growth \textit{per se} is close to worthless, and should therefore not be the primary concern of governments (Andrew Oswald, 1997).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{happiness_vs_income.png}
\caption{Happiness and Real Income Per Capita in the US, 1973-2004}
\end{figure}

\textit{Source}: World Database of Happiness, and Penn World Tables. Real Income data (I$) only available to 2000. Average responses to the following question: ‘Taken all together, how would you say things are these days? Would you say that you are…?’ Responses are coded as (3) Very Happy, (2) Pretty Happy, (1) Not too Happy. Happiness data drawn from the General Social Survey.

This ‘paradox’ is found in US data, but also elsewhere. The same picture can be drawn for Japan (Easterlin, 1995), which has seen one of the largest increases in real per capita income of any country since World War II, and also in Europe. Figure 2 shows trends in average life satisfaction for five European countries since 1973. As in the US, there has been no obvious increase in life
satisfaction over a thirty-year period, even though real incomes per capita have increased sharply in all five countries. The only perceptible trend is in Italy, the poorest country of the five, where average life satisfaction increased from 2.67 in 1973 to 2.88 in 2004, a rise of 9.3%.

FIGURE 2: Life Satisfaction in Five European Countries, 1973-2004

```
0  0.5  1  1.5  2  2.5  3  3.5  4
```

Source: World Database of Happiness. Average responses to the following question: ‘On the whole how satisfied are you with the life you lead’. Responses are coded as (4) Very Satisfied, (3) Fairly Satisfied, (2) Not Very Satisfied, and (1) Not at all Satisfied. Life satisfaction data drawn from the Eurobarometer Survey.

Data from transitional countries, however, tell a different story. Consider Figure 3, which shows average life satisfaction and real income in East Germany during the decade following reunification. East Germans experienced a substantial increase in real income between 1991 and 2002, and reported a considerable rise in their life satisfaction.

This review focuses on the relationship between income and happiness, given its fundamental importance for both economic theory and policy design. The recent growth of the literature on income and happiness has produced a welter of new information. In this paper we will bring these new findings together. In this respect this paper builds upon the general review of the literature on happiness by Bruno Frey and Alois Stutzer (2002a) which mentioned some work on income and happiness published up to 2001. While we note the main findings with respect to income identified by Frey and Stutzer, we will place greater focus on the more recent literature,
especially work using panel data from surveys such as the British Household Panel Survey (BHPS), the German Socio-Economic Panel Study (GSOEP) and the Russian Longitudinal Monitoring Survey (RLMS), which allows researchers to track individuals’ income and happiness over long periods. Informative ‘popular science’ reviews of the general happiness literature can be found in Frey and Stutzer (2002b), Bernard van Praag and Ada Ferrer-i-Carbonell (2004) and Richard Layard (2005).

FIGURE 3: Life Satisfaction and Income in East Germany, 1991-2002

Source: Frijters et al. (2004b). Data drawn from the German Socio-Economic Panel Study (GSOEP). Respondents are asked: ‘How satisfied are you at present with your life, all things considered?’ The responses run from 0 (completely dissatisfied) to 10 (completely satisfied).

The recent increase of interest by economists described above, however, can obscure the fact that psychologists have being putting serious effort into studying the correlates and determinants of various well-being measures for at least the past forty years. Recent comprehensive reviews of this literature include the edited volume by Daniel Kahneman, Ed Diener, and Norbert Schwarz (1999), Ed Diener, Eunkook Suh, Richard Lucas and Heidi Smith (1999), and Richard Lucas, Ed Diener and Randy Larsen (2003). It is clear from these reviews that economists can learn an enormous amount from psychologists about the factors that make people happy. Perhaps the most salient
lessons are that individuals typically adapt to a considerable extent to changes in their circumstances, that people are concerned about their relative standing (e.g. income, status) and not just about their absolute position in society, and that up to 70% of the variation in individuals’ reported well-being can be explained by fixed personality traits. Manipulating these traits is well outside the remit of public policy, so that about one-third of human happiness is potentially amenable by public policy. However, to the extent that some personality traits are developed in early childhood, and come about through nurture rather than nature, this malleable percentage is potentially far higher.

This leads us to the following question: ‘Given that psychologists have been working on this most important of human issues4, why did it take so long for economists to enter the debate?’ Perhaps the main reason is that happiness is hard to measure directly, and most economists are sceptical about any attempt to do so. This contrasts sharply with the indirect traditional economic method of measuring happiness via revealed preferences, where individuals are always assumed to make utility-maximising decisions. The following general questions arise regarding the usefulness of subjective happiness data. First, does asking someone to rate their current level of happiness on some scale really provide a good assessment of their current circumstances (relating to, for example, family life, job and income), or does their answer also reflect ephemeral factors like having a cold or toothache, the weather, their sports team’s recent results, or having recently watched a hard-hitting documentary about third-world famine? Second, can we compare the responses of different people to questions like ‘All things considered, how happy are you with your life?’, to conclude that individual A is happier than individual B? Or even that individual A is happier today than she was a month ago? That is, can we carry out interpersonal comparisons in happiness scores?

This review begins by discussing the (mainly) psychological literature that has addressed these questions. We will argue that there is both a great deal and a wide variety of reliable evidence available suggesting that there is ‘valid variability’ in happiness data (Diener, 1984). This is arguably most strongly demonstrated by the finding that happiness scores today are powerful predictors of future behaviour (for example, quitting one’s job, marriage duration, and life expectancy). The booming economics of happiness literature is testament to the fact that a growing number of economists are convinced that self-reported well-being data contain valuable information. If such measures are good proxies for individual utility, their analysis can provide

---

4 It is often argued that happiness is the ultimate goal of life. Such a goal, for example, is explicit in the United States Declaration of Independence. As noted by Frey and Stutzer (2002), the Declaration ‘takes is as a self-evident truth that the pursuit of happiness is an unalienable right, comparable to life and liberty’ (p. 402).
powerful additional insights into the factors motivating individual behaviour, and therefore the assumptions that underpin economic theory and policy design.

The potential of happiness data to inform contemporary policy debates has already been demonstrated over a wide range of issues. In the realm of macroeconomics, Rafael Di Tella, Robert MacCulloch and Andrew Oswald (2001 and 2003) used cross-country happiness data from Europe to calculate the trade-off between inflation and unemployment, and to show that aggregate happiness is strongly correlated with macroeconomic aggregates. Di Tella and MacCulloch (2005), in the same vein, conclude that there is no difference in this inflation-unemployment trade-off between the rich and poor. Paul Frijters, John Haisken-DeNew and Michael Shields (2004a and 2004b) have demonstrated the large happiness gain to East Germans following Reunification. Paul Frijters, Ingo Geishecker, John Haisken-DeNew and Michael Shields (2005) have shown that the large swings in life satisfaction in post-transition Russia can be partially explained by variations in real incomes. Lastly, Alberto Alesina, Rafael Di Tella and Robert MacCulloch (2004) consider the demand for redistribution between the US and Europe via the effect of income inequality on individual well-being.

In microeconomic analysis, happiness data has been used to better establish the psychological cost and scarring effects of unemployment (e.g. Andrew Clark and Andrew Oswald, 1994; Andrew Clark, Yannis Georgellis and Peter Sanfey, 2001; Liliana Winkelmann and Rainer Winkelmann, 1998). Andrew Clark and Andrew Oswald (1996) showed that workers are sensitive to their ‘relative’ or ‘comparison’ wage, and Alois Stutzer (2004) emphasised the role of income aspirations in explaining happiness. Andrew Clark (2003) and Rafael Laliv (2002) compare the occupational and industry structure of wages and job satisfaction data to conclude that there are wage rents in some sectors of the labour market. Bernard van Praag and Barbara Baarsma (2005) recently used happiness data to estimate the monetary differential necessary to compensate people for the noise of living close to an airport. Johannes Schwarze and Rainer Winkelmann (2005) show that happiness data can be fruitfully applied to the study of interdependent utility, such as the extent of altruism between parents and children. Jon Gruber and Sendhil Mullainathan (2005) relate the well-being of smokers to taxes on cigarettes to conclude that smoking is at least to an extent a bad choice from which smokers would prefer to be rescued (via higher tobacco taxes).

In terms of the subject of this review, there are a number of challenges for economists in establishing the relationship between income and happiness. One key issue is to determine what the relevant measure of income is. This is particularly salient in the case of ‘relative’ income, where we do not know precisely to which incomes individuals compare. These can be ‘external’, coming from others in the same household, in the same neighbourhood, in the same community, or in the
same workplace. Alternatively, they can be ‘internal’, referring to myself in the past or expectations of what I should earn in the future. A second key issue is the difficulty of establishing a causal effect of income on happiness. Just as income, through more consumption, can lead to increased well-being, it is equally likely that happier people have better prospects in the labour or marriage markets. There is also the possibility that both income and happiness are partially jointly determined by unobserved factors (either fixed or time varying), such as personality traits, childhood characteristics or discount rates, which are typically not measured in the large-scale surveys analysed by economists. We have relatively few ‘natural experiments’ in the form of large exogenous shocks to income which can be used to demonstrate causality. This latter point underscores the potential importance of analysing transition countries, which have seen substantial changes in economic, political and social outcomes over a short period of time.

This review is organised as follows. In Section 2 we discuss the key findings of the large literature on the validity of self-reported well-being data. We then provide a coherent theoretical framework for understanding the many facets of the relationship between income and happiness in Section 3. We review the contribution of economists to the income-happiness debate in Section 4, paying particular attention to the most recent literature and to work using household panel surveys which track individuals’ changes in income and happiness over many years. In Section 5 we consider the implications of these findings for key economic issues such as economic growth, consumption, savings, investment, labour supply, taxation, wage profiles and migration. We conclude in Section 6, pointing to a number of gaps in the literature and suggesting topics for future research.

2. Measuring Happiness

It is an important starting point before undertaking any economic analysis of happiness data to know whether individuals mean what they say when they reply to survey questions on well-being. That is, do their answers contain information about what economists call utility? The empirical analysis of happiness is carried out on the understanding that such measures are positive monotonic transformations of an underlying concept called welfare, denoted by $W(.)$, and that happiness is interpersonally ordinally comparable (so that my five, on a one to seven scale, means the same thing as your five). If there is noise in the relationship between $W$ and happiness, then the chances of correctly identifying any relationship between income and happiness is substantially reduced.

Fortunately, there exists a rich empirical literature in Psychology and other disciplines on this topic, which has yielded the broad conclusion that replies to questions on happiness or well-being,
constitute useful and reliable measures of underlying welfare. Some of the different methodological approaches used by Psychologists to investigate this issue are described below.

**Cross-Rater Validity**

The empirical work in this domain is based on the presumption that asking X how happy she is will provide information about her unobserved real level of happiness. A simple validity check is then to ask Y whether he thinks X is happy. Research in Psychology has shown that individuals are able to a large extent to recognise and predict the satisfaction level of others. In interviews in which respondents are shown pictures or videos of others, they demonstrated a high level of accuracy in identifying whether the individual shown to them was happy, sad, jealous, and so on (see Ed Sandvik *et al.*, 1993; or Ed Diener and Rich Lucas, 1999). This was also the case when respondents were shown individuals from other cultures. It is thus arguable that there is a common human 'language' of satisfaction, so that subjective well-being is observable and comparable among individuals. The second finding is that individuals in the same language community have a common understanding of how to translate internal feelings into a number scale, simply in order for individuals to be able to communicate with each other. Respondents have been found to translate verbal labels, such as 'very good' and 'very bad', into roughly the same numerical values (e.g. Bernard Van Praag 1991). It might be argued that there is an evolutionary advantage in being able to gauge how well others are doing.

The general idea of having a third party evaluate respondents’ happiness has been used to validate the replies that individuals themselves provide (REFERENCE NEEDED?). An obvious choice for such a study is individuals who know the respondent well. When friends and family are asked about how happy they believe the respondent is, the scores that they give tend to correlate with the respondent’s own report. Another obvious choice is the interviewer: again, the answer the interviewer gives tallies with that of the respondent. Last, respondents are sometimes given open-

---

5 This is reminiscent of work in the area of ‘emotion’ undertaken in the 1960s (see Chapter 1 of Daniel Nettle, 2005, for a description). American respondents were very good at identifying the emotions depicted by American actors in a series of photographs; but so were the Dani tribespeople of Papua New Guinea.

6 More precisely, it looks as if individuals convert the verbal labels into cardinal numbers that equally divide up the response space. Practically, this is one reason why ordinal and cardinal estimation techniques applied to subjective measures of well-being or health most often produce similar results.

7 This test is not as clean as it might appear at first sight, for the reasons underlined in Charles Manski (1993). Third parties and respondents may share unobserved characteristics which lead them to supply similar answers, even though the correlation between the two underlying constructs (how happy the individual thinks she is, and how happy her friend thinks she is) may be only small. This applies particularly to third-party reports from both friends and family who likely share with the respondent idiosyncratic uses of language. This is less of a problem with third-party raters who are unknown to the respondent. The correlation is not affected if both X and Y use the same linear transformation of their real evaluation into a report, but it will be exaggerated if part of both X’s and Y’s answer reflects a norm belief that people “should” score 8 out of 10 on subjective well-being scales.
ended interviews in conjunction with standard questions about their well-being. When third parties, who do not know the respondent, are played these open-ended interviews their evaluation of the respondent’s well-being matches up with the respondent’s own reply.

**Test-Retest Correlation**

Whereas the other approaches here concern validity (is the measure correlated with the underlying construct of interest?), test-retest correlations are used to evaluate instrument reliability (do you obtain the same outcome when you repeatedly apply the measure?). The reliability of a measure is defined by Randy Larsen and Shane Fredricksen (1999) as the degree to which measures reflect the “true” amount of the construct. As this latter is unobservable in many domains, they note that “For certain psychological constructs, a test-retest correlation is a good estimate of reliability” (p.42). This does suppose, however, that the underlying construct is stable over time (so that any differences found will be between individuals), and it is arguable to what extent well-being measures fall into this category. Test-retest correlations are high when the measures are taken fairly close together, partly because the personality element (the fixed effect) does not change but also because many individuals’ lives change little from one month to the next. It is worth noting that when changes do occur (such as marriage, divorce, unemployment, birth of first child and so on), there are sizeable movements in subjective well-being.

**Physiological**

Instead of relying on others’ judgements, individuals’ well-being replies can be examined in relation to various physiological and neurological phenomena. It is known (see Peter Shizgal, 1999; Fernandez - Dols and Ruiz- Belda, 1995; Sandvik *et al.*, 1993) that there is a strong positive correlation between emotional expressions like smiling, and frowning, and answers to well-being questions. Tiffany Ito and John Cacioppo (1999) showed that positive and negative emotions are associated with the extent of the startle response, and various measures of facial expressions (facial electromyography).

A recent literature has looked at the relationships between positive and negative states, on the one hand, and neurological measures. Obtaining physical measures of brain activity is a huge step in showing that individuals’ self-reports reflect real phenomena.8 Particular interest has been shown

---

8 Richie Davidson (2004) notes that “The identification of the brain circuitry responsible for different aspects of affective processing has helped to parse the domain of emotion into more elementary constituents in a manner similar to that found in cognitive neuroscience, where an appeal to the brain has facilitated the rapid development of theory and data on the subcomponents of various cognitive processes” (p.1395).
in prefrontal brain asymmetry.\(^9\) In right-handed people, positive feelings are generally associated with more alpha power in the left prefrontal cortex, and negative feelings with more alpha power in the right prefrontal cortex.\(^{10}\) This relationship was initially suggested by the observations of patients with unilateral cortical damage (see Davidson, 2004, Section 4). More recently, it has been investigated using techniques to measure localised brain activity such as electrodes on the scalp in Electro-encephalography (EEG) or using scanners in Magnetic Resonance Imaging (MRI).

A recent example is Urry et al. (2004). In this study 84 right-handed individuals (drawn from the Wisconsin Longitudinal Study) provide answers to questions on positive and negative affect, measures of hedonic well-being using global satisfaction with life scores, and measures of eudaimonic well-being, which are designed to capture elements such as mastery, relations with others, self-acceptance and purpose. Brain activity is measured via EEG. The results showed that left-right brain asymmetry is associated with higher levels of positive affect, and with both hedonic and eudaimonic well-being. Interestingly, the correlation between brain asymmetry and positive affect explained all of the correlation with hedonic well-being, but only some of the correlation with eudaimonic well-being. In other words, left-right asymmetry is not just about pleasurable feelings. Davidson (2004) describes further results in which left-right asymmetry is associated with quicker recovery from negative affect challenge, over and above its effect on baseline well-being.

Brain asymmetry is also associated with physiological measures, such as cortisol and corticotropin releasing hormone (CRH), which are involved in response to stress, and with antibody production in response to influenza vaccine (Davidson, 2004, pp.1404-1405). In general, it seems that brain asymmetry is not only associated with measures of SWB, but general measures of wellness of the organism’s functioning.\(^{11}\)

Given the general positive outcomes associated with brain asymmetry, it is important to know why it comes about. It seems likely that there is a role for genetics, with recent work showing that the form of a certain gene involved in regulating the serotonin system (5HTT) is a predictor of neuroticism, which is related to left-right asymmetry (see Nettle, 2005, Chapter 5). However, the degree of asymmetry is far from being a fatality. Davidson (2004, Section 9) underlines the role of early social experiences in determining some aspects of brain circuitry; along similar lines Huppert (2005) talks of the plasticity of the brain in the young. In adults, it is possible to manipulate this

---

\(^9\) Other approaches have also been explored. Brian Knutson et al. (2001) explore the relationship between positive emotions and activity in subcortical circuits including the nucleus accumbens.

\(^{10}\) This is an oversimplification, and recent work has cast the left-right opposition in terms of approach versus withdrawal (anger, a negative approach-related emotion, is associated with more alpha power in the left prefrontal cortex): see Heather Urry et al. (2004).

\(^{11}\) A recent review article by Pressman and Cohen (2005) describes the relationships between affective “style” and health.
left-right balance by showing respondents pleasant or unpleasant pictures or films, and also by using magnetic fields to stimulate activity in the left frontal portion of the brain. Davidson (2004, p.1407) also mentions some intriguing results from a controlled experiment in which those randomly assigned to a meditation group (compared to a neutral control) showed an increase in left-right brain activation; remarkably the meditation group also showed an increase in antibody production in response to influenza vaccine relative to the control group. Although the long-run duration of these effects is not known, it is clear that “the wiring in our brains is not static, nor irrevocably fixed” (Davidson, 2004, p.1407).

**Predicting Outcomes**

Perhaps the evidence that is most persuasive to economists is that respondents seem to act on what they say: they behave as if they were maximising their subjective well-being. In this sense, one obvious test of the hypothesis that subjective data contain useful information about underlying welfare (and therefore about the motives for behaviour) is to use panel data and see if subjective well-being at time $t$ predicts future behaviour.

The empirical analysis has appealed to both probit and duration models to predict future behaviour using panel data. The key result is that, conditional on standard right-hand side variables (such as demographics, income and prices, if relevant), satisfaction or happiness scores do predict behaviour, in the sense that individuals choose to discontinue activities associated with low levels of well-being (see Kahneman et al., 1993; Frijters, 2000; and Shiv and Huber, 2000).

In the labour market, perhaps the most obvious expected correlation is between job satisfaction and quits: workers who are dissatisfied (in cross-section data) should be more likely to quit in the future (if job satisfaction scores can be compared between individuals). Freeman (1978) used US panel data to show that job satisfaction is indeed a strongly significant predictor (often more so than are wages) of quits. This finding has been replicated in other data from the US (see, for example, George Akerlof, Andrew Rose and Janet Yellen, 1988, and McEvoy and Cascio, 1985), Denmark (Nikolai Kristensen and Niels Westergaard-Nielsen, 2004), Germany (Clark et al., 1998) and Great Britain (Clark, 2001). A recent example using data on the self-employed is found in Yannis Georgellis et al. (2005). Further research on the labour market has found that job satisfaction is negatively correlated with absenteeism (Clegg, 1983) and positively correlated with productivity (Patterson et al., 2004); a useful summary is provided in Peter Warr (1999). Clark (2003) shows that mental stress scores on entering unemployment predict unemployment duration: those who suffered the sharpest drop in well-being upon entering unemployment were the quickest to leave it.
The medical literature has also found high correlations in the expected sense between well-being measures and first entry into mental hospitals (Goldberg, 1972), Coronary heart disease (Sales and House, 1971), strokes (Felicia Huppert, 2005) and length of life (Palmore, 1969). Lastly, measures of life satisfaction have also recently been shown to predict future marital break-up (Gardner and Oswald, 2006).

A final point worth making is that when asked in surveys to report their level of happiness, life satisfaction or well-being, it is clear that the vast majority of respondents have no difficulties in provide an answer. For example, less than 1% of respondents in the BHPS or GSOEP do not provide such an answer.

3. Conceptual Framework with Illustrations

In order to help organise our thoughts about the relationship between income and happiness we start by providing a simple conceptual framework, where we specify the following utility function:

\[
U_t = U(u_1(Y_t), u_2(Y_t/Y_t^r), u_3(T - I_t, Z_t))
\]

(1)

where \(U(.)\) is a function common to everyone indicating how the subutilities \(u_1, u_2, \) and \(u_3\) are combined into the final utility \(U(.)\). Here \(U_t\) is (a monotonic transformation of) utility for which reported happiness or life satisfaction is the empirical proxy. We will most often interpret \(U_t\) as the happiness of an individual, but in some indicated instances we will use it for the aggregate happiness of all individuals in a given country. In this specification, \(Y_t\) is the vector of incomes \(y_t\) from \(t=0\) to \(t\) and \(u_i(.)\) can be thought of as the classic consumption function, which is increasing at a decreasing rate in its argument. Given that we speak of a vector of incomes in this general case, this allows for the possibility that incomes in the past may have a positive effect on current consumption i.e. through a wealth effect. In a one-period model or in a model without savings, income will equal consumption \(c\), and \(u_i(Y_t) = u_i(y_t) = u_i(c_t)\).

Here, \(Y_t^r\) is a reference or comparison income, which itself can be made up both of the distribution of income in the entire population, or the income of the same person in the past or in the future, or even the income of individuals in other countries. The relevant set of incomes embedded in \(Y_t^r\) usually determine whether a study looks at a person over time, a person within a family, a person in a firm, a person within a region, a country, or even a whole set of countries. The
function \( u_2(Y,|Y^\ast) \) can be interpreted as the 'status benefit' of income, or alternatively, as the positional aspect of income or the conspicuous consumption aspect of income. The function is presumed to be increasing at a decreasing rate in \( Y \) but decreasing in an increasing rate in \( Y^\ast \). Generically, \( Y^\ast \) will increase in the entire population at the same rate as \( Y \) and in the generic case \( u_2(aY,|aY^\ast) = u_2(Y,|Y^\ast) \) i.e. status is unaffected by a proportional increase of income and reference income simultaneously. In many cases, \( u_2(Y,|Y^\ast) = c \cdot \bar{c} \) where \( \bar{c} \) is average consumption, but the formulation is sufficiently general to encompass the bulk of the specifications used in the literature.

The subutility function \( u_3(T-l,|Z) \) denotes the influence of other factors, whereby the first factor \( (T-l) \) is related to income generation in the sense that it denotes leisure with \( l \) denoting hours at work. In one-period no-wealth models, consumption would be \( lw \) with \( w \) being the wage rate. Finally in this equation, \( Z \) represents a vector of other socio-economic and demographic variables.

**Illustration 1: Consumption and Relative Motives I**

When discussing the empirical literature on the relationship between income and happiness, we will mostly be concerned with very simple re-formulations of (1), such as:

\[
U_i = \beta_1 \sqrt{\ln(y_i)} + \beta_2 \ln(y_i / y_i^\ast) + Z_i y
\]  

(2)

where \( y_i \) is most often a measure of real household income, and \( y_i^\ast \) (if included in the empirical model) is the income of the region, (in rare cases) the explicitly stated reference income, or the income predicted from a wage regression. In order to equate such a function with traditional utility functions which depend on consumption levels, we have to assume the absence of saving.

To see the main forces at work in this representation, take the following stylised implication of what the relationship between income and happiness would look like across countries if i) income is the only systematic difference between countries, and ii) the main aspect of the reference income is the income within a country. In Figure 4, this means that \( U_i = \beta_1 \sqrt{\ln(y_i)} + \beta_2 \ln(y_i / \bar{y}_i) \) where \( \bar{y}_i \) is the average income in the country individual \( i \) resides.\(^{12}\)

\(^{12}\) We here use the subutility function \( \beta_1 \sqrt{\ln(y_i)} \) rather than a term \( \beta_2 \ln(y_i) \) because log income rises almost
The main prediction from this model is that the gradient between income and happiness is steeper within a country, than over time in a country, because within a country there is the status benefit of additional income to the person with greater income. Importantly, however, this status effect has no benefit on the aggregate happiness (in this set-up, the more status one person has, the less all the others have status). Over time in a given country, the main positive effect of income on aggregate happiness is via the consumption component of the utility function.

Illustration 2: Consumption and Relative Motives II

Illustration 1 reflects a pervasive opinion of the last few decades about the relationship between income and happiness at the level of an individual country. It fits the early Easterlin (1974) hypothesis and concurs with much of the psychological and some of the economics literature that we will review later. An additional issue of importance arises when we consider the relationship between income and happiness across several countries simultaneously, such as in the influential study by Di Tella, MacCulloch and Oswald (2003). These authors estimated inter-temporal regressions of year-by-year happiness levels using country fixed effects, time dummies, and macro-variables such as lagged GDP. They found, as Kapteyn et al. (1976) argued, that ’social reference linearly over time, which would imply linearly increasing happiness. The stylised fact that happiness hardly goes up after a certain level of economic wealth thus needs a consumption subutility that is more convex than $\ln(y_t)$. 

The main prediction from this model is that the gradient between income and happiness is steeper within a country, than over time in a country, because within a country there is the status benefit of additional income to the person with greater income. Importantly, however, this status effect has no benefit on the aggregate happiness (in this set-up, the more status one person has, the less all the others have status). Over time in a given country, the main positive effect of income on aggregate happiness is via the consumption component of the utility function.

Illustration 2: Consumption and Relative Motives II

Illustration 1 reflects a pervasive opinion of the last few decades about the relationship between income and happiness at the level of an individual country. It fits the early Easterlin (1974) hypothesis and concurs with much of the psychological and some of the economics literature that we will review later. An additional issue of importance arises when we consider the relationship between income and happiness across several countries simultaneously, such as in the influential study by Di Tella, MacCulloch and Oswald (2003). These authors estimated inter-temporal regressions of year-by-year happiness levels using country fixed effects, time dummies, and macro-variables such as lagged GDP. They found, as Kapteyn et al. (1976) argued, that ’social reference linearly over time, which would imply linearly increasing happiness. The stylised fact that happiness hardly goes up after a certain level of economic wealth thus needs a consumption subutility that is more convex than $\ln(y_t)$.
spaces' can include whole countries, that happiness in one country is very strongly positively related to GDP growth in the last year. To square this finding with the general observation that over long periods of time GDP in richer countries does not appear to be related to happiness, one needs a happiness function such as:

\[
U_{ijt} = \beta_1 \ln(y_{ijt}) + \beta_2 \ln(y_{ijt}^* / y_{jt}^*) + \beta_3 \ln(y_{jt}^* / y_t^*) + Z_{ijt}'\gamma
\]  

(3)

where \( U_{ijt} \) is the happiness of an individual \( i \) in country \( j \) at time \( t \); \( y_{jt}^* \) is average income in country \( j \) at time \( t \); \( y_t^* \) is aggregate income in the set of countries at time \( t \). Note that this happiness function is the same as before within countries, but has an added component \( \beta_3 \ln(y_{jt}^* / y_t^*) \) that reflects the relative income of one country versus a set of countries. We can interpret this component as the ‘warm glow’ effect of feeling part of a successful country or the aggregate degree of optimism.

Importantly, if all countries were to grow at an equal pace, then the factor \( y_{jt}^* / y_t^* \) does not change over time and the interpretation above is simply valid for each country although individual countries at the same time may be on different parts of the curve depending on their level of economic wealth. However, when one country has high GDP growth relative to other countries, then \( y_{jt}^* / y_t^* \) will increase and that country will have a higher degree of happiness that year. This is illustrated by the next 3-dimensional picture shown in Figure 5, where the aggregate happiness in one country depends on the aggregate income in that country and in the rest of the world.

In this case the optimal situation for country \( j \) is to have a high aggregate income while the other countries have low aggregate incomes. Yet, on the central axis, where the income in country \( j \) equals the incomes in the other countries, the additional benefit of more income is subject to decreasing returns. This type of happiness function can be used to help explain why countries are locked in a competition to grow, even though on aggregate that growth only brings benefit via the consumption function: that is for each country separately the component \( \beta_3 \ln(y_{jt}^* / y_t^*) \). However, the happiness benefit from having a higher average income than other countries is from a world perspective is a zero-sum game. Note that if this happiness function is the ‘true’ representation, but we estimate a happiness function without controlling for incomes in the other countries (i.e. the \( y_t^* \)), then the negative effect of the growth in \( y_t^* \) will show up as a downward trend variable, which was found by Di Tella, MacCulloch and Oswald (2003).
Illustration 3: A Status-Race Happiness Function

Another type of happiness function embedded in the general formulation (1) is a function whereby there is a negative feedback from activities that lead to greater individual incomes. Consider for instance:

\[
U_i = \beta_1 \ln(y_i) + \beta_2 \ln(y_i / y_j) + \gamma \ln(T - y_i / w_i)
\]  

(4)

where the main difference with the previous examples is the component \( \gamma \ln(T - y_i / w_i) \) which can be interpreted as the happiness derived from leisure time, where leisure time equals a maximum \( T \) minus the number of hours spent on earning income level \( y_i \) in the case when wage levels equal \( w_i \). Consider Figure 6 that gives an illustration of what the actual choice would be relative to the happiness optimising choice:
Figure 6: The Marginal Happiness of Additional Income for an Individual versus a Country

where the top curve shows $\frac{\partial U_i}{\partial y_i}$, which is the marginal utility to the individual of obtaining additional income via additional time worked; the lower curve represents $\frac{\partial U_i}{\partial y_i} + \frac{\partial U_y}{\partial y_i} | y_i = y''$ which is the effect of additional income in the country in the case where each individual makes the same choice. We can see that the second curve lies below the first one because of the negative externality of $y_i$ in the component $\beta_2 \ln(y_i/y_t)$. Individuals will choose the point $Y2$ where their marginal utility of additional income is zero, whereas in fact it would be optimal for society if individuals choose $Y1$.

Illustration 4: An Adaptive Happiness Function

A final illustration concerns the possibility that happiness is adaptive in the sense that individuals get used to their circumstances and that changes have only transient effects. The notion of adaptation can be seen as the result of a backward looking reference position. Consider for instance the following specification:
where the main difference with the previous examples is that the reference income level $y^*_t$ is now entirely made up of backward looking elements of the income stream of the individual herself.

Under the presumed specification of the reference income, the final equation to be estimated (third line) consists only of the logs of current income and the income in the last 3 periods. This equation can in principle be extended to include further lags of current income and can even include expected future incomes. The main characteristics of this specification is that the short-run effect of an increase in log income equals $\beta_1 + \beta_2$ whilst the long-run effect is only $\beta_1$.

\[ y^*_t = (y^*_{t-1})^\alpha (y^*_{t-2})^\gamma (y^*_{t-3})^{1-\alpha-\gamma} \]

\[ U_{yt} = \beta_1 \ln(y^*_t) + \beta_2 [\ln(y^*_t) - \alpha \ln(y^*_{t-1}) - \gamma \ln(y^*_{t-2}) - (1 - \alpha - \gamma) \ln(y^*_{t-3})] \]

**4. Implications.**

**Introduction**

Many of the implications of the empirical investigations on income and happiness can be thought of as informing us about the following simple utility representation:

\[ U_t = U_1(c_t) + U_2(y_t | y^*_t) + U_3(T - I) \]

The left-hand side here denotes the utility of an individual at time $t$, and for now denotes both the utility of choices, as well as the anticipated outcome of choices: in the language of Kahneman and Tversky (1997, QJE), it refers to both experienced utility and decision utility. The first function on the right-hand side denotes utility from personal consumption ($=c_t$) where we can think of the sheer enjoyment of being able to sustain a healthy lifestyle. The second function denotes the utility of relative income, where the income of an individual at time $t$ ($=y_t$) depends on some reference income point ($=y^*_t$). The third function denotes a non-materialistic term, i.e. leisure. It depends on total hours available minus the hours spent at work. We’d think of each of the sub-utility functions as increasing in their argument, but subject to decreasing marginal utility. We’d also think that in a life-cycle context, personal consumption $c_t$ is linked to personal incomes $y_t$ via a budget constraint, and that reference incomes on average grow at the same rate as personal incomes. In a static model, income and consumption are the same. There are many variations of (1) one can think of, such as
dropping the assumption that utility is linearly separable or that there is a reference income point rather than a reference consumption point, but for presentational purposes it is handy to think of utility being made up of some absolute material term, some relative material term, and some non-material term. One can give a simple evolutionary rationale for why utility functions would have this kind of generic form: the term $U_1(c_t)$ ensures maximum effort to escape dire poverty in the case that basic survival is threatened (i.e. when $c_t$ is very low), thus ensuring the survival of the individual; the term $U_2(y_t | y^*_t)$ ensures an individual, and then especially a male, is competing relative to others however advantageous the current overall situation is, which ensures relative (procreation) effort remains high in all circumstances. This argument was made forcefully by Coelho and McClure (1993). Samuelson (2004) has an explicit evolutionary model where continual changes in the general environment make a relative utility function evolutionary preferable to an absolute utility function; Finally, the term $U_3(T - l)$ can either be seen as ensuring some form of social investment like actually caring for children or maintaining social relations that are ultimately useful for procreation, or as a function that ensures an individual is not in ‘over-drive’ all of the time, i.e. as a function guaranteeing that individuals now and then take a much-needed break. We can already foreshadow the main policy implications by noting that an increase in an individual’s personal income $y_t$ has a negative externality on the happiness of other individuals via an increase in the overall reference income (= $y^*_t$). Frank (1985) and Layard (2005) argue forcefully that this implies over-investments in activities related to getting higher income, at the expense of ‘unconspicuous’ (or ‘nonpositional’) activities such as caring for children and social relations in general. Both Frank (1985) and Layard (2005) advocate higher taxation on income on this basis. Dupor and Liu (2003) point out that the case for social intervention depends on the existence of both some reference point that affects the marginal utility of own consumption, which they term ‘keeping up with the Joneses’, and the existence of activities that can be crowded out because of this. In terms of (1) the essential point of Dupor and Liu (2003) is that if there was no term $U_3(T - l)$ then there would be no immediate reason to tax income because nothing is crowded out. Only intertemporal issues to do with when to consume how much then remain. We will in a subsequent section go more deeply into these issues when we address the theory literature on relative incomes and optimal taxation.

4.1 Economic growth.
We can place several of the arguments made in the literature in terms of this simple equation. First, take the debate about whether economic growth leads to greater happiness or not. Easterlin (1974, 1995) argues that economic growth in Western nations does not lead to greater happiness, witness the fact that happiness levels have not significantly moved in Western countries over time. Yet, in countries that start out from a much poorer level, such as East Germany (Frijters et al. 2004b) and Russia (Frijters et al. 2005a) income growth has been associated with modest increases in happiness. In terms of the model above, Western countries seem to be at a point of personal consumption $c_t$ where the marginal utility from $U_1(c_t)$ is minimal, whilst for poorer countries there is still a gain in $U_1(c_t)$ from additional personal consumption. Further economic growth then has no noticeable aggregate effect because on aggregate, reference incomes move up with the same pace as incomes, so there is no net gain in $U_2(c_t')$ to a country as a whole from additional income increases. This is presuming for the moment that economic growth has no effect on the distribution of income and thus only affects consumption levels. The argument above is then exactly the argument that Layard (2005), Clark and Oswald (2002), McBride (2001), and Veenhoven (1999) also make about economic growth and happiness. Equivalent formulations are the assertion that at certain levels of development, only conspicuous consumption is important, or that ‘keeping up with the Joneses’ is the main economic motive in rich countries.

A related issue comes in when we relax the assumption that economic growth and the distribution of income are unrelated. Then, the effect of inequality on aggregate happiness becomes relevant too. If we think of the second sub-utility function $U_2(c_t')$ as convex and everyone in a country sharing the same reference income (i.e. some national norm), then it’s immediate that a country with the same level of aggregate incomes, personal consumption, and work choices with more unequal income distributions would be less happy on average. This is an additional rationale for income equality over and above the usual argument that consumption equality has welfare benefits because of convexity in the subutility $U_1(c_t)$. The effect of economic growth on happiness then depends on whether economic growth is related to more or less inequality.

There is a flip side to the argument that greater economic prosperity at some point ceases to buy more happiness. One can argue that it is the concern for relative incomes embodied in the second sub-utility function $U_2(c_t')$ that keeps economic growth going after some wealth level. The argument is that relative concerns are more important in rich countries, because the importance of personal consumption has become a marginal concern in rich countries. Hence, what makes individuals work hard in rich countries despite high aggregate consumption levels is the concern for
status. This is indeed a possible evolutionary reason for the term $U_2(y_t \mid y^*_t)$ being there in the first place. This argument has a long ancestry in economic debates. Bernard Mandeville’s 1705 ‘Fable of the Bees’ puts the argument allegorically. Mandeville juxtaposes two hypothetical beehives: one in which the bees only care about sustaining themselves and have no interest in status (i.e. there is only $U_1(c_t)$ and $U_1(T - l)$) and another beehive where relative status is what mainly matters (i.e. there is mainly $U_2(y_t \mid y^*_t)$). Mandeville postulates that the first beehive would be happy, but not very rich and ultimately doomed to be taken over by the second beehive where the bees are mainly motivated by status (by $U_2(y_t \mid y^*_t)$). In that second beehive, Mandeville argues, the bees would keep working and keep looking for opportunities both within and outside their beehive to further their relative standing, leading to continual expansion and growth of the second beehive. Mandeville’s observations have been echoed by many others since. Adam Smith, in his Theory of Moral Sentiments, put it thus: ‘To what purpose is all the toil and bustle of this world?... It is our vanity which urges us on.’. The modern-day equivalent of the argument by Mandeville and Smith and many other early economists is made in theoretical models by Glomm and Ravikumar (1994), Corneo and Jeanne (2001), and Pham (2005). The key aspect of those models is that they specify $U_1(c_t)$ to be $\ln(c_t)$ and $U_2(y_t \mid y^*_t)$ to be $\ln(k_t) - \theta \ln(k^*_t)$ where $k_t$ denotes wealth in stead of income. Those models abstract from the possibility of leisure. In their models too, it is the $U_2(y_t \mid y^*_t)$ part of the utility function that drives continued economic growth. Their argument is related to that made by Fershtman, Murphy and Weiss (1996) who argue that social status is determined in some part by relative education which thus links economic growth via education to status considerations. Normatively speaking, the dominance of the status motive in the income/happiness gradient means the benefits of economic growth are not to be found in happiness increases. There are other reasons indirectly related to utility that would still give a rationale for economic growth, much in the vein of Mandeville’s arguments: the link between the length of life and (aggregate) income; the link between the ability to withstand foreign aggression and levels of economic activity; the ability to attract migrants when income levels are relatively high; and some status utility benefit to a country as a whole from having high levels of income compared to other countries. Each of these possibilities relates to other literatures we will only touch upon in the remainder.

**4.2 Labour supply**
Mandeville and his successors directly predict labour supply to remain high during economic expansion because of status motives. Neumark and Postlewaite (1998) similarly note that in models where only personal consumption matters and there is decreasing marginal utility to consumption, one would invariably expect aggregate labour supply to go down with aggregate consumption levels, just as the bees in Mandeville’s first beehive cease to work hard. In the above utility function (1) however, there is a limit to the long-run reduction in labour supply with increased consumption levels, because the relative considerations in $U_2(y_i \mid y_i^*)$ are undiminished with increased consumption levels. Neumark and Postlewaite argue that this status aspect of the income/happiness link is the main explanation for why levels of work participation have not declined dramatically in the 20th century even though consumption levels have increased dramatically. The same conclusions arise if we’d see a job, rather than the income associated to it, as the vehicle of status: then too, labour supply is rather unresponsive to overall consumption levels. Woittiez and Kapteyn (1998) and Vendrik (1998) point out that there may be intermediary factors between relative income concerns and labour supply decisions, such as social norms with respect to appropriate labour supply decisions that themselves in the long run are determined by the payoffs to a more basic utility function. Woittiez and Kapteyn (1998) and Vendrik (1998) argue that female labour supply has increased in many countries because of changing wage profiles (i.e. a change in the non-work/work utility trade-off), but much slower than one would expect because the associated social norms took time to adjust.

4.3 Poverty

A direct implication from the stylised finding that much of the income/happiness gradient is made up of relative considerations relates to the measurement of poverty. Theoretically, we’d want to think of poverty as a state of low utility, though we usually mean some narrow, material, notion of utility. To make clear how we distinguish between utility and whatever we usually mean by poverty, think of a factor like sunshine. No known definition of poverty takes it as relevant whether a materially poor person enjoys more sunshine than a rich person suffering in a cold climate, even if the materially poorer person is happier. Implicitly, sunshine and other non-material factors influencing happiness are not counted as relevant to poverty, even though it is both relevant for happiness and for individual decision making. Poverty rather concerns the $U_1(c_i)$ and $U_2(y_i \mid y_i^*)$ parts of utility, in stead of happiness proper. One can distinguish the absolute and relative poverty concepts by whether they take account of the $U_2(y_i \mid y_i^*)$ part of the utility function or not. The whole notion of basic needs and absolute poverty is based on the idea that poverty means being
below a certain minimum level of $U_1(c_i)$, whatever the level of reference incomes. This is for instance the notion of poverty used in official US statistics. The notion of relative poverty championed by Sen (1976) and many others explicitly recognises that individual compare themselves to others and that poverty is thus mainly about low relative standing, such as having a lower income than the half-mean or half-median, which turn out to be the most commonly used poverty measures internationally.

### 4.4 Migration.

When we consider more deeply what the elements are of the reference position $y_i^*$, several other implications of the income/happiness relation come out.

Consider the decision to migrate or not. If there was no term $U_2(y_i \mid y_i^*)$ then all individuals who could earn more in another country would move there according to the simple utility representation in (1), abstracting from sunk costs. This is not true if we start specifying $y_i^*$ however. For instance, if $y_i^*$ equals the average income in the neighbourhood one lives or the average income of those like yourself at your workplace, then a person who rightfully fears ending up at the bottom of the relative income distribution in another country would not migrate to that country even if the absolute income she could earn there is higher. Stark and Taylor (1991) use this idea to rationalise why the elites of poor countries don’t emigrate: those elite are at the top of the income distribution in the country they live and would not be so if they emigrate. Stark and Taylor furthermore introduce the notion of different economic migrant types. Put in terms of (1), they argue that one would expect that those with relatively high earning potential in another country, i.e. those whose skills are relatively undervalued in the country of origin, would emigrate. Furthermore, one would expect that those who keep comparing themselves to individuals in their home (poorer) country would actually be more likely to emigrate than those who adapt their reference income to the host country: they can become a high status person in the home country by earning more in the host country yet nevertheless in a status sense keep being a part of the home country. This rationalises why many migrants keep visiting their home country: it’s when they cash in on their high status as relatively high incomes earners compared to those in the home country, making it irrelevant whether those incomes are high relative to the host country. A final implication of this line of thinking is that there are two distinct possible migration dynamics. In the first scenario, those who migrate do so voluntarily because of the $U_2(y_i \mid y_i^*)$ part of their utility function: they have exceptionally high skills and can become high-status in a host country and want to compare
themselves to persons in the host country anyway. Hence the first group of voluntary migrants would be made up of high-status assimilators and Stark (2004, WD) identifies them with brain-drain phenomena. The networks of these early migrants can lead to more migration of the same variety but because of the fast assimilation of such migrants this is not something that particularly occupies those migrants meaning their networks will not be solely dominated by other migrants from the same home country. The second migration pattern is very different. A second type of migration wave is that lead by exceptionally low consumptions back home (i.e. by the $U_t(c_t)$ part).

One can think of the Irish in the US driven by the potato famine back home, forced migration in general, or the ‘guest worker phenomenon’ of the European union where whole villages were essentially transplanted to other countries in the 60’s. Those individuals will by design be less likely to assimilate and, indeed, will have strong incentives to keep comparing themselves to people in the home country because of their high wages relative to those in the home country and low wages relative to the host country. Those migrants will for the same reason attempt to attract more low-skill individuals from the home country, because those do not detract from their own status but indeed increase it by decreasing the average reference income they experience in the host country. The two types of immigration waves, associated with the two different parts of the income/happiness gradient, thus have very different cultural and economic implications for the host country. Countries like Australia and Canada who operate a points system whereby a potential migrant has to offer something exceptional to the host country in order to attain a visa, attempt to attract the first type of economic migrant. Countries in the EU that are introducing legislation trying to stop ‘family reunions’, i.e. marriages between migrants and individuals in the home country, can be interpreted to be trying to stop the second kind of economic migrant.

4.5 Optimal taxation and consumption.

There has been a real burgeoning in recent years of the optimal tax implications of reference incomes, which all rely on and lead to predictions about the precise empirical properties of the income/happiness relations. This theory literature has not in the past been directly connected to the empirical happiness literature, but, as we will see, there are potential gains for those two literatures to become more integrated (as, for instance, Weinzierl 2005 advocates).

An early paper looking at optimal taxation was by Frank (1985), whose utility function read:

$$U = U(c_0, R(c_0), c_1)$$

(1a)
Here, $c_0$ is the consumption of some positional good, which potentially includes both positional material goods (conspicuous consumption) and positional immaterial goods (conspicuous leisure).

The first term of Frank’s utility function corresponds to $U_1(c_0)$ in (1). The second term in the utility function $R(c_0)$ denotes the rank of the consumption of a positional good of a person in the distribution of the consumption of that good. This term corresponds loosely to $U_2(y_t | y_t^*)$ in (1).

The third term $c_1$ denotes a nonpositional good and corresponds loosely to $U_3(T - l)$ in (1).

Frank’s basic point is that individual maximisation will mean they consume $c_0$ until the point where total marginal utility is zero:

$$\frac{dU}{dc_0} |^{R(c_0)} + \frac{dR(c_0)}{dc_0} \frac{dU}{dR(c_0)} |^{R(c_0)} = -\frac{dU}{dc_1} U(c_0, R(c_0), c_1)$$

where $\frac{dU}{dc_0} |^{R(c_0)}$ is the marginal utility of the consumption of the positional good keeping the rank constant, $\frac{dR(c_0)}{dc_0} \frac{dU}{dR(c_0)} |^{R(c_0)}$ is the marginal utility of the consumption of the positional good via its effect on rank, and $\frac{dU}{dc_1} U(c_0, R(c_0), c_1)$ is the indirect effect of the increased consumption of the positional good via the (reduced) consumption of the nonpositional good. The precise form of $\frac{dc_1}{dc_0}$ is given by the budget constraint that fixes total income, allowing for nonpositional goods $c_1$ to include both leisure activities and consumption activities. Frank’s main point is that $\frac{dR(c_0)}{dc_0} \frac{dU}{dR(c_0)} |^{R(c_0)}$ is a pure externality because changes in rank have no social benefit even though they have a private benefit. This additional benefit of positional goods to an individual leads to societal overconsumption of positional goods to the detriment of non-positional goods. Frank pointed out that this externality gives a rationale for taxation on the positional good in order to promote the nonpositional good. If we equate the positional good as relative income and the nonpositional good as leisure, one would get a rationale for income taxation in order to promote leisure. Layard (2005) adopts this argument entirely.

Other authors take a different specification of the utility function and of the reference position in particular, leading to different tax implications.

Ireland (1994) takes the following utility function:

$$U = U(f(c_0, c_1), s(c_0))$$
where \( c_i \) is a good whose consumption is unobservable, and \( s(c_0) \) is status, specified as the belief spectators have about \( f(c_0, c_i) \) based on observing \( c_0 \). Exactly like above, Ireland shows a general tendency for overconsumption of the observable good \( c_o \) for a wide class of possible information regimes. Ireland shows the Pareto improvements that can be attained with income taxation coupled with direct transfers of \( c_i \) to the poor.

Ljungqvist and Uhlig (2000) look at a similar utility function but are interested in changes in optimal tax policy over the business cycle. Their main utility function is:

\[
U = \frac{(c_i - \alpha \overline{c}_i)^{1/\gamma} - 1}{1 - \gamma} - \beta l
\]

with \( \overline{c}_i \) the average of the population of the individual consumption \( c_i \) and \( l \) labour supply. Note that this utility function has a similar set-up as (1), albeit that \( c_i \) and \( \overline{c}_i \) don’t enter in separable functions. The externality embedded in the presence of \( \overline{c}_i \) leads to labour supply that is too high, very much in the same mould as Boskin and Sheshinski (1978) and Frank (1985). They show this externality can be perfectly countered by a constant marginal tax on \( c_i \), independent of the business cycle. Ljungqvist and Uhlig extend their analysis by looking at the case when the reference position is not \( \overline{c}_i \) but \( c_{i-1} \), i.e. average consumption last period. In the presence of aggregate productivity shocks, they find optimal tax rates then co-move with current productivity shocks which thus create counter-cyclical effects of taxation on the economy, exactly in line with the usual Keynesian optimal tax policy which is also counter-cyclical.

### 4.6 Savings and investment in the presence of relative concerns.

A recent theory literature has emerged on the dynamic effects of relative consumption concerns (status concerns). The papers we look at abstract from the possibility of a term like \( U_3(T - l) \) and look mainly at the timing of consumption. The optimality of the savings decision is the main question.

One could naively think that status concerns have no effect on savings, because the trade-off between status now and status in the future would seem to be the same as the trade-off between consumption now and consumption in the future. This naïve thought breaks down when one considers that individuals accumulate wealth over their lifetime and productivity generally increases, implying that in a stylised sense the reference income when old is ‘always’ higher than when young. Then, relative concerns change the marginal utility of consumption within different
periods of the lifecycle and thus affect saving and investment decisions. Most of the generic arguments that come out of this line of thinking can be seen by closely examining the utility function introduced by Abel (1990) and subsequently adopted by many others:

$$U_t = \frac{(c_t)^{(1-\eta)(1-\alpha)}}{1-\alpha} \cdot \left( \frac{c_t}{\bar{c}_t} \right)^{\alpha(1-\alpha)}$$

where $c_t$ is own consumption and $\bar{c}_t$ is the geometric mean of the consumption of a reference group, which could be the rest of the population or some slowly adjusting social norm containing all previous generations, or even prior consumption of the person involved. Here $0<\eta<1$ denotes a kind of ‘weight’ of relative concerns to the individual; and $\alpha >0$ affects risk-aversion ($\alpha =0$ denotes risk-neutrality). The important characteristics of this utility function come out when we consider the marginal individual utility of own consumption when the reference income $\bar{c}_t$ is high or low for various choices of $\alpha$. The main possibilities are shows in the graph below:

The two lines with $0<\alpha<1$ show how a higher reference income reduces own utility, but also reduces the marginal utility of own consumption. This implies that when $0<\alpha<1$, individuals will want to have high consumption levels when the reference consumption is low. This creates a coordination problem in the case that the reference income is due to choices of other persons, such as when the reference income is the average of the consumption of others. Then, individuals will want to plan consumption when other individuals are not consuming. The difficulty in solving this coordination problem in endogenous growth models has so far appeared unsurpassable (Ljungqvist and Uhlig (2000) and Abel (2005) presume for their main results that $\alpha >1$). Yet, it intuitively seems plausible to presume that the marginal (status) utility of consuming is higher when the
consumption of others is lower. Intuitively also, one would thus think that when $\alpha < 1$, savings is too low because individuals will not postpone consumption till the future because consumption levels in the future will generally be higher due to productivity growth.

Exactly the opposite reasoning appears for $\alpha > 1$, which corresponds to the lower utility profiles of the graph above. Then, again, the utility of the individual is lower with higher reference incomes, but the marginal utility of consumption is higher. Then, individuals want to consume more when others consume more, implying a kind of herding phenomenon takes place: status is then a bandwagon good in the terminology of Duessenberry (1949). Then, there is no problematic coordination issue because all consumers will want to consume at the same time and the simple situation arises that all individuals save too much at the same time because they all want to consume as much as possible at the end of their lives when the consumption of others is higher too due to productivity increases.

Importantly, the issue of whether $\alpha > 1$ can in principle be empirically examined in the happiness literature, though this has not to our knowledge happened so far. One way to examine the sign of $(\alpha - 1)$ becomes clear if we consider cross-derivatives of $(1d)$:

$$\frac{d^2 U_t}{dc_t dc_{\tilde{c}_t}} = \eta (\alpha - 1) (c_t)^{1-\alpha} \tilde{c}_t^{-\eta(1-\alpha)-1}$$

from which follows that the sign of the interaction of own consumption and the reference consumption is the same as the sign of $(\alpha - 1)$. This can be the basis of an identification strategy.

Abel (2005) focuses on this issue in an overlapping two-generations model by using a utility function for the new-born of the form

$$U_t = \frac{(c_t)^{(1-\eta)(1-\alpha)} * (c_t / \tilde{c}_t)^{\eta(1-\alpha)}}{1-\alpha} + \beta (c_{t+1})^{(1-\eta)(1-\alpha)} * (c_{t+1} / \tilde{c}_{t+1})^{\eta(1-\alpha)}$$

where $c_t$ is own consumption at time $t$ and $\tilde{c}_t$ is the geometric mean of the consumption of all others alive at $t$. The term $(c_t)^{(1-\eta)(1-\alpha)}$ is the part of own consumption independent of the consumption of others, and the term with $\beta$ denotes the importance of future consumption. The essential point emerges when we look at the marginal effect of an increase in future consumption:

$$\frac{dU_t}{dc_{t+1}} = \beta (c_{t+1})^{1-\alpha} \tilde{c}_{t+1}^{-\eta(1-\alpha)}$$

when $\alpha < 1$ one gets the intuitive result discussed above, which is that the marginal return to status consumption in the future is lower when $\tilde{c}_{t+1}$ is higher and thus when there are general production increases. This externality of higher consumption in the future reduces savings and provides a case for subsidies on savings. Abel discusses cases when $\alpha > 1$ under which general productivity
increases leading to higher future reference incomes \( \tilde{c}_{t+1} \) actually increase the future marginal return to status, yielding a case for taxes on savings (capital taxes). Carrol, Overland, and Weil (1997, 2000) make a different point by using a utility function where the reference position is solely determined by the consumption of the individual herself in the past:

\[
U_t = \left( \frac{c_t}{\tilde{c}_t} \right)^{1-\eta} \left( \frac{c_t}{c_t} \right)^{\rho(1-\alpha)}
\]

\[
\tilde{c}_t = \int_{-\infty}^{t} e^{-\rho(s-t)} c_s \, ds
\]

which is the same utility function as Abel (1990, 2005) but with a reference position that is purely self-determined. These authors also only consider \( \alpha > 1 \) and thus indeed generically argue there is upward pressure from technological growth on savings because individuals anticipate the marginal return to future status to be higher. However, there is in the specification of the reference position \( \tilde{c}_t \) another effect related to the speed with which individuals’ reference position adjusts to current consumption. When \( \rho \) is high, adjustment of the reference position is high, and when \( \rho \) is small, though always positive, adjustment is slow and consumption of a long time ago still matters. Note that the speed of adjustment can in principle be measured in happiness regressions. Carrol, Overland, and Weil (2000) show that when adjustment is slow, it makes sense for individuals to save more when economic growth is faster despite the fact that future reference consumptions will be higher under higher economic growth rates. The intuition is that under slow adjustment, an individual wants to ‘smooth out income increases’ more than under fast adjustment. When the reference position adjusts very fast, one essentially wants to enjoy the status benefit of a productivity gain immediately. In this way the authors rationalise their empirical claim that countries with high growth rates also have high savings rates, which is an empirical regularity that cannot be easily reconciled with a model without reference positions affecting utility: without reference positions, the marginal utility of future consumption is always lower under faster economic growth because of the higher level of future consumption, thus leading to lower savings rates. Why save now if one is rich tomorrow anyway? The answer of Carrol, Overland, and Weil (2000) is that one wants to adjust slowly to ever high consumption patterns and it thus makes sense to save more now, simply to avoid getting used to high consumption patterns too soon.

Constantinides (1990), and Galli (1994) use a similar utility function to address the equity premium puzzle. Constantinides notes that the existence of an adjusting self-reflective consumption level gives an additional reason for individuals to only slowly adjust their consumption over time. Deaton (1992) confirms this prediction by showing for US data that individuals excessively smooth of consumption after shocks to permanent incomes. Constantinides (1990) goes further than this by
arguing that the presence of $\tilde{c}_t$ creates a bias in conventional estimates of risk-aversion that essentially look at the tradeoffs people make over time: without the presence of $\tilde{c}_t$ one would think rich individuals would save much more than poor individuals under high-risk aversion. The finding that there is not much variation across income groups in saving rates can then lead one to think risk-aversion can’t be that high. Constantinides (1990) shows that the presence of $\tilde{c}_t$ breaks this logic in that risk-aversion can nevertheless be very high, but that the presence of a moving reference position leads all individuals to smooth income over time in a similar manner. This way, Constantinides rationalises instances of high instantaneous risk aversion, in particular the large premium individuals are prepared to pay for risk-free assets versus higher-return assets with risk (the equity premium puzzle). Cambell and Cochrane (1999) extend this idea to more types of asset pricing puzzles and argue that a model including a reference position $\tilde{c}_t$ has superior predictive power to models without such a term.

4.7 Wage profiles.
Kahneman (1999) found in choice experiments that individuals are, at the margin, about twice as sensitive about losses as they are about gains which he terms loss-aversion or status quo bias. In terms of our simple utility presentation, the reference aspect of the utility function $U_2(y_t | y_t^*)$ would then have the following kind of form:

$$U_2(y_t | y_t^*) = 2f(y_t - y_t^*) - f(| y_t - y_t^* |) * I_{[y_t > y_t^*]}$$

$y_t^* = \sum_{s=0}^{t} w_s y_{t-s}$

where the term $f(| y_t - y_t^* |) * I_{[y_t > y_t^*]}$ denotes the additional penalty of having incomes below the reference position and the reference income is some weighted average of previous incomes. The graph below shows what such a utility function would look like in the case that the reference income equals income in the last period, and compares an increasing income profile with a decreasing income profile:
This utility function has the property that losses weigh twice as heavy as gains, at least locally. In the graph, the reference income next period is closely related to income this period, as many authors propose. Note how the increasing and decreasing income profiles have been drawn to perfectly overlap for the three periods depicted: the individual gets the same undiscounted income in both profiles, but gets them in increasing order in the top profile and in decreasing order in the bottom profile. This leads to much lower utility levels in the decreasing profile compared to the increasing profile, because the person getting the increasing profile is forever getting more than his reference income whereas the person getting a decreasing profile is forever getting less than his reference income. Yet, the actuarial value of the decreasing income profile is higher: any positive discount rate means getting the higher income first has a higher present-day value than getting the lower incomes first. This means an employer can make a profit by offering an increasing profile with lower actuarial value than a decreasing one with a higher actuarial value because this is what individuals would choose even if they are perfectly rational. Frank and Hutchens (1993) and Loewenstein and Sicherman (1991) indeed showed this choice behaviour to hold in psychological choice experiments (Andrew: ???).

The strong probability for the existence of such a kink in the utility version, also termed loss aversion, has more implications than just to rationalise upward sloping tenure profiles within firms. It also means one would expect employers to offer contracts guaranteeing no income reductions during a working life, i.e. an endogenous absence of demotion and an institutionalised downwards wage rigidity. A working paper by McDonald (2002) motivates an asymmetric utility function with
this ‘loss-aversion’ finding, and then argues in a simple micro-macro model that it creates downward wage rigidity at the level of firms which in turn generates Keynesian business cycles. It fits the empirical observation of Teulings and Hartog (1997) that wage decreases are virtually never observed within organisations because individuals get sacked rather than demoted.

A potential avenue of future research along these lines is the hypothesis that retirement is strongly related to individuals being unwilling to take a step back within their organisation and thus choosing to retire rather than accept a wage cut or a demotion. Retirement would then generically follow the moment at which the productivity of an individual peaks, even though an individual could still have many highly productive years left, simply because individuals are loath to accept a job and an income that is below their current reference position. In such a situation, there would be a case for deferring rewards for production till later in life, i.e. to smooth out wages such that they increase till some fixed age, which in turn raises the issue of credible long-term contracts and mandatory retirement.

4.8 The ability of individuals to anticipate changes in reference incomes: the case for paternalism.

The experiments of Frank and Hutchens (1993) and Loewenstein and Sicherman (1991) also show that individuals are somewhat conscious of the fact that their reference positions will adapt over time and make career choices consistent with that expectation. Indeed, the entire theory literature reviewed above presumes rational expectations with respect to reference positions. Still, many other authors have argued that individuals are particularly poor in anticipating that they reference position adapts over time. Easterlin (2001) argues individuals for that reason overestimate how happy they’ll be in the future.

The empirical laboratory regularities of an ‘endowment effect’ and a ‘status quo bias’, which are both based on the fact that individuals resist changes more ex ante than ex post, nevertheless indicate that individuals to at least some degree do fail to anticipate they’d make different tradeoffs in the future. There are two alternative interpretations of this kind of finding. One is that individuals are simply inherently incapable of anticipating they’ll see things differently in the future. It’s an unintentional mistake that people would rectify if made fully aware of. The second interpretation is that individuals only care about the future insofar as the reference position then overlaps with their reference position now. They literally see a future self as only relevant if that future self has a similar reference position. That’s not a mistake, but rather like hyperbolic discounting: it’s inherent in the way a rational Self 0 evaluates the future. The basic paternalistic question arising from this possibility has been well put by many, including Kahneman et al. (1997, 1999), Rabin (1998), and
Simon: should we care only for what a Self 0 wants or should we care about what a Self 1 experiences. If one cares only about the wishes of Self 0 then one chooses to take decision utility as the normative reference point. If one takes the experiences of Self 1 as the guiding principle of policy, this means potentially going against the explicit wishes of an ‘ignorant’ electorate. Many individuals point out that government policies capture both possibilities: via education and state propaganda we actively try to ‘civilise’ our population in the belief that those educated and patriotic are better people after their socialisation than before. Yet, via the electoral process we from a certain age onwards take the desires of the Self 0’s as the guiding motivation for government policy though even then many government policies have only long-term effects on selves that have no presence now, like environmental policies affecting future generations.

Whilst this debate is primarily philosophical in nature, the literature on the income/happiness has lead to explicit empirical evidence for the idea that individuals indeed do take into account that they will over time adapt their reference positions to their own experiences, although they take the changes in the incomes of others much less into account. The normative implication would be indeed that this leads to an under appreciation by individuals of the negative externality of other people’s income increase on the status of others and thus to an aggregate over-emphasis on economic growth. Making this argument, Layard (2005) advocates growth-reducing taxation designed to stimulate greater degrees of leisure. We may note here that one can take the model in (1) one step further by the recognition that status games not only involve income, but also time-investments. Veblen (1898) recognised this by talking about conspicuous consumption and ‘conspicuous leisure’, which consists of all time investments whose main payoff is showing other people that one can afford to spend time on leisure such as via useless skills (such as speaking latin, and playing a musical instrument, which each signal an abundance of time not used to increase $U_i(c_i)$). The tax implication would then be mixed because one would ideally want to tax all investments into status races equally in order to promote non conspicuous consumption and leisure. Layard (1980) even went so far as to recognise that one may ideally want to sustain several status races rather than fewer because more races have more winners. The question then arises what sustains multiple status races, whilst still countering any crowding-out effect on non-status goods and leisure activities. Frijters and Leigh (2005) hypothesise that conspicuous leisure diminishes in mobile environments because mobility reduces the visibility of leisure more than of consumption. They indeed find that US states with higher rates of mobility have higher aggregate levels of investment in conspicuous consumption (i.e. higher labour supply), both at any moment in time and through time. The average hours worked per week in 1981-2003 in US states with the highest level
of internal mobility is about 28, versus 26 hours in a standard week for states with the lowest level of internal mobility. Frijters and Leigh advocate mobility taxes in order to restore the balance between conspicuous consumption and conspicuous leisure.

4.9 As-if models.

We may here review some situations in which a status concern can arise as a reduced form property of a model with an underlying utility function lacking any status element. One possibility is the argument made by Cole, Mailath and Postlewaite (1992) who argue that a concern for relative status is an endogenous feature of the marriage market. In their marriage market individual’s of two different types (i.e. genders) match and a person with a relatively high value within their type are more likely to match with a person with a relatively high value from the another type. Even though the basic utility function has only a \( U_i(c_i) \) (in their case, a CRRA utility function) and some stake in bequests, the reduced form utility function of men includes the wealth rank in one of their equilibria. The essential reason for this is that because the other gender is maximising, its rank relative to others of the same gender is in the indirect utility function. Another possibility is discussed in the classic tournament model of Lazear where a manager can engender optimal effort out of a group of employees by promising to give a large prize to the person perceived to have worked the hardest. In such a world, the individual pay-back function is of the form in (1)\(^\text{13}\) and there is indeed a negative externality on one worker’s utility from the efforts of others to win the tournament. Yet, the increased general effort is not suboptimal for overall utility because the increased effort of others also increases the size of prize in equilibrium, which is a positive externality perfectly off-setting the negative one. The behavioural implication at the individual level of tournaments is the same as that of a status-seeking utility specification implying that tournaments provide an as-if rationale for using (1) without making aggregate normative statements. There are other circumstances in which it may be optimal for a principle to set individuals a relative reward function. The generic situation is where the agents have more information about aggregate productivity than principles because then a contract encouraging the different agents to compete, thereby revealing their shared information, has benefits to the principle over standard piece-rate contracts. This kind of situation arises in financial markets where fund-managers are paid more when they perform better than the average fund manager; it arises in benchmark regulation of privatised utilities; it arises in winner-take-all markets like court cases and

\(^{13}\) If we interpret \( y_i \) as total individual observed production and \( y_i^* \) as the highest observed production of the others, then \( U_2(y_i \mid y_i^*) = P^* \frac{I_{y_i > y_i^*}}{P} \) with \( P \) the size of the prize and.
sports matches. In all such cases, individual pay-back has a large relative component to it yet this
does not imply aggregate welfare to be unrelated to aggregate incomes. Yet all those cases provide
examples of how the individual implications of (1) carry over in an as-if way to many fields of
economic activity. Indeed, Samuelson (2004) argues that the competitive effect of relative
consumption concerns were what lead ‘Nature’ to build relative concerns into the utility function in
the first place.

4.10 Difficulties
We should not underestimate the challenges that still remain for empirical results on the
happiness/income relationship to inform mainstream theory literature. These difficulties remain
even if income and happiness are perfectly measured and conform to experienced utility and
purchasing power. To name but a couple of remaining problems:
- Theory often dictates that the relevant measure is consumption, not income. The difference
  between the two is savings and a closer correspondence between the happiness literature and the
  theory literature would need better estimates of savings. Savings are notoriously hard to measure
  empirically in panels. Not just do individuals have difficulty remembering their bank savings when
  asked, but more fundamentally is it difficult to empirically establish the borderline between a
  purchase that has only current consumption benefits and a purchase with some future consumption
  benefit. How much of a car or a house purchased today should be counted as current consumption?
  How much is education a form of current status consumption and how much is it an investment?
  Such issues, which relate to the majority of our purchasing decisions, are extremely tricky and
  make for a large potential gulf between theoretical models of savings and empirical estimates of
  savings.
- To measure reference groups is very difficult, especially when individuals move a lot in their
  lifetimes and reside in high-population-density areas. The ‘best’ attempt we know is by Knight and
  Song (2004) who were not just able to identify the Chinese villages their respondents came from,
  but were able to confirm that 90% of individuals indeed saw their village as their reference group
  by simply asking them whom they compared themselves to, making their rural sample of happiness
  very suited to address the question of how important reference groups really are. Melenberg
  1985???. Most of the rest of the literature has had to guess reference incomes and thus inserts things
  like the predicted income of the characteristics of an individual or the income in the whole state,
  which is much less convincing. The issues become even harder when one considers the fact that
  urban individuals can to some extend choose their reference group (Falk and Knell, 2004).
- The precise behaviour of happiness when own income is very close to the reference income can only be addressed when one has very precise data on both the timing of income changes and the timing of reference income changes. Panel data interviewing individuals once a year is basically inadequate to pick up the asymmetry in happiness to changes that push income below the reference income versus changes that push it above. For instance, consumption changes may precede changes in permanent incomes. Individuals may hear about a future promotion, which is a permanent income increase known to happen in advance which could increase consumption at the moment of the information shock, whilst observed income only moves later. This implies the happiness increase precedes the observed income increase. The reverse may also hold: big-ticket items like children may only be deemed affordable after prolonged increases in income, meaning consumption increases follow income increases with a large lag. Then, the happiness increase follows the income increase. Only with extremely precise data on the timing of income shocks, and the timing of the arrival of information on these income shocks, can one address the asymmetry between the happiness effect of an increase in income above the reference position versus a decrease below. At present, only experiments can address that asymmetry. Experiments have limitations though because subjects in experiments are usually non-representative; because the laboratory situation itself lacks realism; and because laboratory experiments on social phenomena are inherently unsuitable to pick up long-run adaptation (like the adaptation of reference groups) because humans can’t be kept in laboratories for year.

5. Conclusions and discussions.

The stylised findings of the income/happiness literature are that:

1. Income matters at the individual level.
2. Relative income is more important for higher income persons than for lower income persons.
3. Average income growth in rich countries is not related to aggregate happiness increases, but income growth in poor countries does lead to some aggregate happiness growth.
4. Individuals are very sensitive to income reductions relative to others in the same surroundings (village, firm, neighbourhood).

These empirical findings have been used to rationalise many phenomenon. (3) has been used to argue the lack of normative justification behind economic growth; (2) and (4) have been used to argue the importance of taxes on visible consumption; (4) has been used to argue the relevance of wage rigidities leading to Keynesian business cycles; (1) and (2) have been used to argue for
relative poverty lines versus absolute poverty lines; the endogeneity of the reference group in (4) has been used to argue that there are various types of economic migrant, one intent to identify with the new country and one intent to identify with the home country; the interaction between reference consumption and the marginal utility of own consumption has been argued to determine whether savings on a macro-level are too high or too low; the very existence of reference income effects in (2) has been long argued (by Mandeville, Smith, Keynes, and Layard) to be the main reason for humanities’ continued push towards greater prosperity.

The precise relation between income and happiness is thus central to much of micro and macro economics. It’s the core of saving’s decisions, wage policies, and migration decisions. Yet, the empirical literature still has a long way to go before it is precise enough to go beyond ‘informing theory’ towards ‘interacting with theory’. Many of these challenges are shared with other empirical literatures. We still lack the ability to precisely time changes in income with changes in happiness over long time periods, nor are we quite able to precisely map incomes into current and expected consumption. We lack the ability in most of our data to reliably know who is comparing themselves to whom, both in a spatial and a social sense. We lack the ability to adequately deal with the many endogeneities in our data, such as the endogeneity of marriage, education, and family background. Yet, the income/happiness literature is a key aspect of a push within economics towards greater behavioural realism, where we become ever more realistic in our models and empirics about how people really feel and behave.


