The relationship between happiness, health and socio-economic factors: results based on Swedish micro data

Ulf-G Gerdtham* and Magnus Johannesson,

Department of Economics, Stockholm School of Economics, Stockholm, Sweden

Working Paper Series in Economics and Finance No. 207

November 1997

Abstract: This paper investigates the relationship between happiness (utility) and a host of socio-economic variables. The data set consists of a random sample of over 5,000 individuals from the Swedish adult population. Happiness is measured by a three-point categorical measure of overall happiness (not happy, happy sometimes, happy most of the time), and an ordered probit model is used to econometrically estimate the happiness equation. The results are consistent with the theoretical predictions and show that happiness increases with income and education and decreases with unemployment, urbanisation, being single, and male gender. The relationship between age and happiness is U-shaped, with happiness being lowest in the age-group 45-64.

JEL classification: D60, I31, I12

Key words: Happiness, utility, health, socioeconomic factors.

Acknowlegements: We are grateful to Professor William H. Greene for valuable comments on the econometric methods. The financial support from the National Corporation of Swedish Pharmacies is gratefully acknowledged.

*Corresponding author:
Department of Economics
Stockholm School of Economics
Box 6501
S-113 83 Stockholm
Sweden

Phone: +46 8 7369283; Fax: +46 8 302115; E-mail: heug@hhs.se

1. Introduction

The concept of utility has a long tradition in economics. The origin of the concept dates back to the writings of Bentham in the sixteenth and seventeenth centuries (Stark 1952). Bentham referred to utility as a measure of happiness or pleasure and argued that utility was a measurable cardinal quantity that could be compared between individuals. This view of utility prevailed until the beginning of this century, when Pareto (1909) showed that demand theory could be derived based only on information about the ranking of different alternatives. This led to the ordinal utility concept, where utility simply refers to a preference ordering of alternatives and utility cannot be compared between individuals. This ordinal view of utility became the dominant view, and is still held by most economists.

The realization that interpersonal comparisons are necessary for normative issues of economic policy has, however, led to renewed interest in Bentham s concept of measurable and interpersonally comparable utility (van Praag 1971,1991; Christensen et al 1975; Jorgensen et al 1980; Tinbergen 1991; Ellingsen 1994; Clark & Oswald 1994; Saunders 1996; Kahneman et al 1997). Most of the work on estimating cardinal utility functions for interpersonal comparisons has focused on the impact of income and consumption goods on utility (van Praag 1971,1991; Christensen et al 1975; Jorgensen et al 1980; Saunders 1996). Tinbergen (1991), however, argued that happiness is not only a function of consumption, but other factors such as social relations and education may be as important. There is thus a need for studies which examine a broader set of determinants of happiness.[1]

From an equity perspective it may be especially important to assess the relationship between various socio-economic factors and happiness. In a recent article Clark & Oswald (1994) also

[1] The terms ”utility” and ”happiness” are used interchangeably in the paper to represent utility.

assessed the relationship between socioeconomic factors and happiness, with a special focus on the impact of unemployment. They found that unemployment reduced happiness, and that unemployment led to a greater loss in utility than divorce or marital separation. They also found a U-shaped relationship between age and happiness, with unhappiness being lowest in a person s mid-thirties. Surprisingly, they also found a negative relationship between happiness and education and they found no consistent relationship between income and happiness.

A potential weakness in the study by Clark & Oswald (1994) is the measure of happiness used. It is based on twelve questions about psychological distress, with the answers being added together to give a disutility scale between 0 and 12. This scale is then used as an ordered scale for (un)happiness. However, since the dimensions of happiness captured by each of these questions may not be equally important, the scale may not lead to the same ordering as the utility function of an individual.

The aim of this study is to assess the relationship between happiness and a host of socio-economic variables. We measure happiness by a three-point categorical measure of overall happiness (not happy, happy sometimes, happy most of the time). This measure avoids the problem of the measure used by Clark & Oswald. Health status is a factor that can be expected to be an especially important determinant of happiness. Since a number of socio-economic variables may be important for both health status and happiness, we estimate an equation system that makes it possible to estimate both the direct effect of a variable on happiness and the indirect effect through its impact on health status. In Section 2 below a simple theoretical model that is the basis of the empirical analysis is presented. In Section 3 the data and variables used are described and in Section 4 the estimation methods are outlined. Section 5 reports the results and the paper ends with some concluding remarks in Section 6.

2. The theoretical model

Individuals are assumed to derive utility according to the following utility function:

$$U_i=U_i(h_i,x_i,S_i) \qquad (1)$$

where U_i is the utility of individual i (i=1,,,,,,I), h_i is the health status of individual i and x_i is a vector of private goods consumed by individual i. S_i is a vector of socioeconomic factors that affect utility. The utility function is assumed to be cardinally measurable up to a positive affine transformation; in other words transformations of the form a+bU, where a and b are constants (b>0), are allowed. We also assume that the utility function is fully comparable between individuals, i.e. in positive affine transformations a and b are the same for all individuals (Boadway & Bruce 1984). Health is produced according to the following health production function:

$$h_i=f(m_i,h_{0i},S_i) \qquad (2)$$

where m_i is a vector of health goods such as medical care and h_{0i} is the initial (given) health status. Utility is maximized subject to the following budget constraint:

$$y_i=P_ix_i+p_im_i \qquad (3)$$

where y_i is the exogenously given income of individual i, P_i is a vector of private goods prices faced by individual i and p_i is a vector of health goods prices faced by individual i. This leads to the following indirect utility function:

$$V_i=V_i(h_i(y_i,P_i,p_i,h_{0i},S_i),y_i,P_i,S_i) \quad (4)$$

As can be seen in equation (4), the exogenous variables (y,P,S) may influence utility either directly or indirectly through the intervening health variable. This suggests two alternative approaches for estimating the utility model. The first approach is to model the intervening health variable explicitly in the following equation system:

$$V_i= \beta_1+\beta_2 y_i+\beta_3 P_i+\beta_4 h_i+\beta_5 S_i+\varepsilon_1 \quad (5a)$$

$$h_i= \beta_6+\beta_7 y_i+\beta_8 P_i+\beta_9 p_i+\beta_{10} h_{0i}+\beta_{11} S_i+\varepsilon_2 \quad (5b)$$

Alternatively, the health variable can be omitted from eq. (5) and the following reduced form model estimated:

$$V_i=\beta_{12}+\beta_{13} y_i+\beta_{14} P_i+\beta_{15} p_i+\beta_{16} h_{0i}+\beta_{17} S_i +\varepsilon_3 \quad (6)$$

In the above equations β_1-β_{17} are coefficients to be estimated and ε_1-ε_3 are error terms with zero mean and constant variance, and we also assume that cov($\varepsilon_{1,}$ ε_2) is zero. The full structural approach of equation (5) distinguishes between the indirect effects of the exogenous variables working through health and the direct effects of the exogenous variables after controlling for health, i.e. it identifies the process underlying the effects of the exogenous variables. The second approach captures only the total (direct and indirect) effects of the exogeneous variables in a reduced form model. In this paper we use both approaches to evaluate both the direct and indirect effects of the variables. The data used to estimate equations 5 and 6 are detailed below, together with the hypotheses for the different variables.

3. Data and hypotheses

The empirical analysis of the utility model is based on data from a probability sample of the Swedish population, the Level of Living Survey (LNU) from 1991 (Institutet för Social Forskning 1992). The total sample consists of 6,773 individuals, between the ages 18-76 years. After correcting for missing values, the sample was reduced to 5,106 individuals for the total sample. The survey contains data on for instance morbidity and different socioeconomic variables. Further details about the LNU are given in: Institutet för Social Forskning (1992) and Fritzell and Lundberg (1994).

In Table 1 the variables used in the regression analysis are defined and in Table 2 summary statistics for the variables are given. Since there is no variation in the prices of private goods (P_i) and health goods (p_i) between individuals, it was not possible to include these variables in the analysis.

TABLES 1 AND 2 IN HERE

3.1 Utility (V)

Utility (happiness) is measured by a categorical question about life-satisfaction included in the LNU. In this question the individual s rated their personal satisfaction on a three-point scale (0=the daily life is never a source of personal satisfaction, 1=the daily life is sometimes a source of personal satisfaction, 2=the daily life is a source of personal satisfaction most of the time.) The question provides three levels of utility, and due to the assumption of full comparability of utility functions, the utility of the three utility levels will be the same for all individuals. The ordered utility variable is used as the dependent variable in the estimation of the structural and reduced form utility functions.

3.2 Health status (h)

The health status is also measured by a categorical measure. In the categorical health rating question the individual s rated their own current health status on a three-point scale (0=bad health, 1=fair health, 2=good health). This type of categorical health measure has been shown to capture important information about the individual s health (Connelly et al. 1989) and to be an important predictor of mortality (Wannamethee & Shaper 1991, Kaplan & Camacho 1983; Idler & Kasl 1991). This variable is used as the dependent variable in the estimation of the determinants of health equation. It is also used as an independent variable in the estimation of the structural utility equation, by entering two dummy variables for fair health (HEALTH1) and good health (HEALTH2). It is expected that the utility function will be increasing in health status, and we thus expect a positive sign for this variable.

3.3 Income (y)

We use dummy variables rather than a continuous income variable to avoid making assumptions about the functional relationship between income and utility. Gross annual income is measured by three dummy variables: INC2=1 if the income corresponds to the second quartile (61,500-116,300), INC3=1 if the income corresponds to the third quartile (116,300-161,700) and INC4=1 if the income corresonds to the fourth quartile (>161,700). The reference category is individuals who belong to the first (poorest) quartile of the gross income distribution (<61,500).[2] The source of the income data is the National Income Tax Statistics, linked to the LNU data.

The income is given in Skr (Swedish Kronor; exchange rate September 1997 $1=Skr 7.80). Income is expected to have a positive direct effect on utility, and a positive effect on health status. The total effect on utility is thus also expected to be positive.

3.4 The initial health status (h_0)

We use two proxy variables for the initial inherited stock of health capital. We use a dummy variable for overweight measured as a body mass index over 30 (BMI>30), and we also include a dummy variable for if the parents or siblings of the respondent had any health problems (HPROBFMS). Both these variables are hypothesized to decrease the health status, and thus have an indirect negative effect on utility.

3.5 Socioeconomic factors (S)

We include the following socioeconomic variables in the analysis: age, gender, education, civil status, unemployment and urbanisation. Rather than impose a functional form on the relationship between health and age, we conservatively used three 0-1 dummies for age groups (AGE2, AGE3 and AGE4). We expect the health status to decrease with age, whereas it is difficult to know a priori what the direct effect of age on utility will be. The total effect of age on utility is thus indeterminate on the basis of theory.

Gender is represented by a 0-1 dummy for male (MALE). We expect men to have a lower health status than women, due to their lower life-expectancy. We have no prespecified hypothesis about the direct effet of gender on utility and the total effect is thus also indeterminate on the basis of theory. We include two dummies for the education of the individual (EDUC2, EDUC3).

---

[2] The income quartiles are based on the total sample of the 6,773 individuals, which explains why the

We expect education to have a positive direct effect on utility. Since higher educated persons can also be expected to be more efficient producers of health (Grossman 1972; Wagstaff 1986,1993), education is also expected to have a positive effect on health status. The total effect of education on utility is thus also expected to be positive.

As a measure of civil status we include a dummy variable for if the individual is not married or cohabitant (SINGLE). It is expected that being single will have a negative direct effect on utility and a negative effect on health status. The total effect on utility is thus also expected to be negative. We also include a dummy variable for if the individual is currently unemployed (UNEMPL). Unemployment is expected to have a direct negative effect on utility and a negative effect on health status. For the total effect on utility we thus also expect a negative effect.

Finally, we include two dummy variables for urbanisation (SMALLCITY, BIGCITY). We expect urbanisation to have a negative effect on health status, whereas we have no a priori hypothesis about the direct effect of utility on urbanisation. The total effect of urbanisation on utility is thus indeterminate on the basis of theory.

4. Estimation issues

In estimating the models we have to take into account the fact that both our dependent variables (utility and health status) are ordered responses with three categories each. An appropriate tool for analyzing such ordered categorical data is the ordered probit model (for references see Amemiya (1981), Cameron & Trivedi (1986), Greene (1993)). Let V_i^* be a continuous, latent variable representing for instance the cardinal utility function of the individual. We assume a linear dependence between the latent variable V_i^* and X_i, β and ε_i:

---

corresponding mean figures in Table 2 are not equal to 0.25.

$V_i^* = \beta X_i + \varepsilon_i, \varepsilon_i \sim N(0,\sigma^2)$,

The variable V_i^* defines a variable v_i which is related to the above mentioned categories in the following way:

$$v_i = \begin{cases} 0 \text{ if } V_i^* \leq \theta_0 \\ 1 \text{ if } \theta_0 < V_i^* \leq \theta_1 \\ 2 \text{ if } \theta_1 < V_i^* \end{cases}$$

where θ_i = 0,1, are unobservable thresholds. Denoting the cumulative density function of the standard normal distribution as above (Φ), it follows that the probabilities of an individual for each category are given by:

$$\text{Prob}[h_i=0] = \Phi[\mu_0 - \alpha X],$$

$$\text{Prob}[h_i=1] = \Phi[\mu_1 - \alpha X] - \Phi[\mu_0 - \alpha X],$$

$$\text{Prob}[h_i=2] = 1- \Phi[\mu_1 - \alpha X],$$

with $\alpha = \beta/\sigma$ and $\theta_j/\sigma = \mu_{j-1}$, 0,1, i.e. note that only the ratios β/σ and θ_j/σ are estimable (Dustman 1996). If the regression contains a constant term, the full set of coefficients is not identified. A common normalization is to set $\mu_0= 0$, which means that the estimated coefficients μ_i, i=1 represent the differences in the respective thresholds: $\mu_i = \mu_i - \mu_{i-1}$ (Greene 1995; Dustman 1996). Greene (1993) points out that the interpretation of the estimates is not straightforward. A positive regression coefficient indicates that an increase in the respective variable shifts weight from category 0 into category 2, which means that the probability of category 2 increases and the probability of category 0 decreases.

The estimated regression equation, α x, is a positive affine transformation of the underlying variable (where the constants used in the transformation are unknown). This means that in this case with a cardinal utility function as the underlying variable the estimated regression equation is a positive affine transformation of the underlying utility function (V_i*). The estimated regression equation, α x, is thus also a cardinally measurable utility function (unique up to a positive affine transformation), and the estimated regression coefficients can be interpreted as the impact of the variables on cardinal utility.

The ordered probit model was used to estimate the equation system in equation (5), and the reduced form model in equation (6). Equation (5) is a recursive (triangular) system with a diagonal Σ matrix, i.e. there is a unidirectional dependency among health and utility and the disturbances across equations are assumed to be contemporaneously uncorrelated. Both equations in the utility-health system can thus be estimated separately by the ordered probit model (see Greene 1993).[3]

5. Results

In Table 3 we report the estimation results for the structural and reduced form happiness models. To facilitate the interpretation of the results we also show in Table 4 the predicted probability of being in the highest happiness category (happy most of the time) for each level of

[3] If the assumption of zero contemporaneous correlation is wrong we should resort to the Zellner estimation technique, but for ordered probit models. Unfortunately, such a procedure is not available in LIMDEP [6.0], as is the software used for estimation in this study. However, we estimated the two equations as a linear SURE model estimated by two-step feasible GLS, and found that this did not change the qualitative results concerning the directions of the estimated effects of the variables in the equations. These results are available from the authors upon request.

the explanatory variable, at the mean level of all other explanatory variables. The goodness of fit values (pseudo R^2) are between 0.04 to 0.07 depending on the choice of pseudo R^2 in the structural utility equation and somewhat lower in the reduced form model. The goodness of fit value for the health equation in the structural model is about 0.10.

TABLES 3 AND 4 IN HERE

5.1 The effect of health status (h)

As expected, health status has a significant positive effect on happiness. Both the health status dummy variables are significant ($P<0.01$), and the effect of HEALTH2 (good health) is also significantly higher than the effect of HEALTH1 (fair health;$P<0.01$). The predicted probability of being happy most of the time is 0.42 with a bad health status and 0.60 with a good health status.

5.2 The effect of Income (y)

The estimated effects of income on happiness are positive both in the structural and the reduced form models. In the structural model, the estimated direct effects of INC2-4 on happiness are all positive and INC2 and INC4 are significant ($P<0.10$) and INC3 are close to the limit of significance at the 10% level. The estimated effect of INC4 tend also to be higher than INC2-3. As expected, income is also positively related to health status, which means that the total effect of income on happiness will exceed the direct effect. This is also seen in the reduced form model where all the income dummy variables are positive and significant ($P<0.10$) and the effect of income increases successively for each income variable. The predicted probability of being happy most of the time increases from 0.53 in the lowest income quartile (INC1) to 0.61 in the highest income quartile, if the total effect of income is taken into account. If only the direct effect of income on utility is taken into account the difference in the probability of being happy most of the time is only three percentage units between the lowest and the highest income quartiles. This indicates that a large part of the effect of income on happiness takes place through the intervening health variable.

5.3 The effect of the initial health status (h_0)

As expected, our two proxy variables for the initial inherited health status, overweight (BMI>30) and health problems in the family (HPROBFMS), are significantly negatively related to health status in the structural form equation system. Contrary to our hypothesis, however, they have no impact on happiness in the reduced form model.

5.4 The effect of socioeconomic factors (S)

The direct effect of age on happiness is positive and AGE4 (>65 years) is significantly positive, indicating that individuals in the highest age group are happier than individuals in the youngest age group (18-34 years) if health status is held constant. The effect of age on health status is as expected negative and all three age dummy variables are highly significant. This effect on health status will thus counteract the positive direct effect of age on happiness. The total effect of age on happiness is shown in the reduced form model, and in this equation AGE2 (35-44 years) and AGE3 (45-64 years) have negative signs, although only AGE3 is significant (P>0.10). AGE4 (>65 years) has a positive effect that is not significant compared to the baseline category, but is close to significance compared to AGE3 (P=0.1336). Overall the relationship between age and happiness is thus U-shaped, with happiness being lowest in the age-group 45-64 years. The predicted probability of being happy most of the time is 0.59 in the age-group 18-34 years, 0.57

in the age-group 35-44 years, 0.55 in the 45-64 years age-group and 0.60 in the age-group over 65 years.

Male gender has a significantly negative direct effect on happiness. As expected, male gender also has a negative effect on health status, which will further increase the negative effect on happiness. The total effect of male gender is also negative and highly significant in the reduced form model. The predicted probability of being happy most of the time is 0.55 for men and 0.60 for women in the reduced form model.

Education, as expected, has a positive direct effect on happiness, and both the education dummy variables are significant ($P<0.05$). Education also has a highly significant positive effect on health status ($P<0.01$), indicating that education increases the productivity of producing health. This effect will increase the total effect of education on happiness. Both education variables are also highly significant in the reduced form model ($P<0.01$). The predicted probability of being happy most of the time increases from 0.55 with less than high school education to 0.67 for university education, based on the total effect in the reduced form model. This can be compared to an increase in the probability of being happy most of the time from 0.57 (less than high school education) to 0.61 (university education), if only the direct effect on happiness is taken into account. According to the size of the regression coefficients the difference in utility between having a university education and having less than high school education is greater than the difference in utility between men and women ($P<0.05$), but not significantly different from the difference in utility between the highest and the lowest income quartile.

Being single, as expected, has a significant negative direct effect on happiness ($P<0.01$) and a significant negative effect on health status ($P<0.01$). The total effect is thus also negative and highly significant in the reduced form model ($P<0.01$). The predicted probability of being happy most of the time is 0.47 for persons who are single and 0.62 for persons who are married or cohabiting. The effect on utility of being single is greater than the difference in utility between the highest and the lowest income quartile (P<0.05) and the difference in utility between men and women (P<0.01), but is not significantly different from the difference in utility between the highest and the lowest education category.

Unemployment has a direct significant negative effect on happiness (P<0.01), but has no significant effect on health status. The total effect of unemployment is negative and highly significant in the reduced form model (P<0.01). The probability of being happy most of the time is 0.49 for individuals who are unemployed and 0.58 for individuals who are not unemployed. The effect of unemployment on utility is not significantly different from the effect of gender or the effect of being single. The effect is also not significantly different from the difference in utility between the highest and the lowest income quartile.

The estimated effect of urbanisation shows that living in big cities has a significant negative direct effect on happiness (P<0.10) and a significant negative effect on health status (P<0.05), compared to living in rural areas. The total effect on happiness of living in big cities is also negative and significant in the reduced form model (P<0.05). There is no significant difference between living in rural areas and living in smaller cities in the structural form model or the reduced form model. The predicted probability of being happy most of the time is 0.55 in big cities and 0.58 in small cities and rural areas. The effect of living in big cities on utility is significantly lower than the effect of being single (P<0.01), but is not significantly different from the effect of gender or the effect of unemployment.

6. Concluding remarks

We have investigated the relationship between happiness and various socio-economic variables, based on data from a random sample of over 5,000 individuals in the Swedish adult population. Using a categorical measure of overall happiness and regression analysis with the ordered probit model, we are able to estimate the effect on happiness of a host of socioeconomic variables. Both a structural form and a reduced form model are estimated, to distinguish between the direct effect on utility and the indirect effect through the intervening health status variable.

Our results are consistent with the theoretical predictions. As expected, happiness increases with both education and income, and decreases with being single and unemployed. We also find that women are happier than men and that people living in rural areas or small cities are happier than people living in big cities. As regards age, we find a U-shaped relationship between age and happiness, with happiness being lowest in the age-group 45-64 years. Our results also show the importance of taking into account the fact that many variables affect utility both directly and indirectly through the effect on health status. The indirect effect was shown to be very important for the effect of variables such as income, education and being single.

It is interesting to compare our results with the recent study by Clark & Oswald (1994). Our results are consistent with the results of Clark & Oswald in that unemployment and being single has a strong negative effect on happiness. A U-shaped relationship between age and happiness is also found in both studies, with happiness being lowest in the mid thirties in the Clark & Oswald study and in the age-group 45-64 years in our study. There are also some important differences in the results of the two studies. Clark & Oswald found no systematic relationship between income and happiness and found a negative relationship between education and happiness. As predicted by economic theory, we find a positive relationship between both income and happiness and between education and happiness. Clark & Oswald also find that men are happier than women, whereas we find the opposite result. There are many possible explanations for these divergencies in results. They could reflect genuine differences between the two countries or the two samples, or they could reflect differences in the measurement of the different variables. When it comes to the measurement of variables a potential problem in the Clark & Oswald study is that their measure of happiness may not order alternatives the same way as the utility function. This is because the measure of happiness was constructed by adding different dimensions of psychological distress that may not be equally important. That problem was avoided in this study by using a categorical measure of overall happiness.

Our results support the arguments made by Tinbergen (1991), that many non-economic variables may be as important for happiness as income and consumption. Our results suggest that socio-economic variables such as unemployment, health status, gender, being single, urbanisation and education may be as important as income for happiness. This has important implications in for instance studies of equity and distributional issues, since to assess the distribution of welfare in society it is important to also take into account the distribution of non-economic factors such as education, health status and employment possibilities. A further investigation of the relationship between non-economic factors on happiness and the interaction between economic and non-economic variables is a challenging issue for further research.

In conclusion, it is important to note some limitations of the study. One limitation is that there may be omitted variables that affect happiness, which may lead to omitted variable bias. Another important limitation concerns causality. Since we use cross-section data, it is necessary to be cautious in inferring causality between variables. For some variables there could be problems with reverse causality, e.g. that unhappiness or health status affect income and education rather than the other way round. If this is the case, then the estimated effects of income and education will be biased, along with the effects of all other correlated regressors. Unfortunately, the lack of instruments precludes some formal tests of possible endogeneity and our results therefore need to be interpreted with caution.

References

Amemiya, T., 1981, Qualitative response model: a survey, Journal of Econometric Literature 19, 481-536.

Boadway, R.W. and Bruce, N., 1984, Welfare economics, Blackwell, Oxford.

Cameron, C. and Trivedi, P., 1986, Econometric models based on count data: comparisons and applications of some estimators and tests, Journal of Econometrics 1, 29-53.

Clark, A.E. and Oswald, A.J., 1994, Unhappiness and unemployment, Economic Journal 104, 648-659.

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

Connelly, J.E., Philbrick, J.T., Smith, R., Kaiser, D.L. and Wymer, A., 1989, Health perceptions of primary care patients and the influence on health care utilization. Supplement to Medical Care 27, 99-109.

Dustman, C., 1996, The social assimilation of immigrants, Journal of Population Economics 9, 37-54.

Ellingsen, T., 1994, Cardinal utility: a history of hedonimetry, in Allais, M. and Hagen, O., (Eds.), Cardinalism: a fundamental approach, Kluwer Academic Publishers, Dordrecht.

Fritzell, J., Lundberg, O. (red.), (1994), Vardagens villkor. Brombergs. (The Conditions of Daily Living).

Greene, W., 1993, Econometric Analysis, Second Ed., Macmillan Publishing Company, New York.

Greene, W., 1995, LIMDEP™ Version 7.0 User s manual, Econometric Software, Inc.

Grossman, M., 1972, On the concept of health capital and the demand for health, Journal of Political Economy 80, 223-255.

Institutet för Social Forskning (1992), 1991 års levnadsnivåundersökning. Instruktions- och övningsformulär. Stockholms Universitet. (Level of Living Survey 1991).

Idler, E.L. and Kasl, S., 1991, Health perceptions and survival: do global evaluations of health status really predict mortality?, Journal of Gerontology 46, 555-565.

Johansson, P-O., 1995, Evaluating health risks: an economic approach, Cambridge University Press, Cambridge.

Jorgenson, D.W, Lau, L.J. and Stoker, T.M., 1980, Welfare comparison under exact aggregation, American Economic Review 70, 268-272.

Kahneman, D., Wakker, P.P. and Sarin, R., 1997, Back to Bentham? Explorations of experienced utility, Quarterly Journal of Economics 112, 375-405.

Kaplan, G.A. and Camacho, T., 1983, Perceived health and mortality: a nine-year follow-up of the human population laboratory cohort, American Journal of Epidemiology 117, 292-304.

Magee, L., 1990, R^2 measures based on Wald and likelihood ratio joint significance tests, American Statistician 44, 250-253.

von Neumann, J. and Morgenstern, O., 1947, Theory of games and economic behavior, Princeton University Press, Princeton, NJ.

Pareto, V., 1909, Manuel d èconomie politique, V. Giard & E. Brière, Paris.

Saunders, P., 1996, Income, health and happiness, Australian Economic Review 4th quarter, 353-366.

Stark, W., (ed.), 1952, Jeremy Bentham s economic writings, vol. 1-3. George Allen & Unwin, London.

Tinbergen, J., 1991, On the measurement of welfare, Journal of Econometrics 50, 7-13.

Van Praag, B.M.S., 1971, The welfare function of income in Belgium: an empirical investigation, European Economic Review 2, 337-369.

Van Praag, B.M.S., 1991, Ordinal and cardinal utility: an integration of the two dimensions of the welfare concept, Journal of Econometrics 50, 69-89.

Wagstaff, A., 1986, The demand for health: some new empirical evidence, Journal of Health Economics 5, 195-233.

Wagstaff, A., 1993, The demand for health: an empirical reformulation of the Grossman model, Health Economics 2, 189-198.

Wannamethee, G. and Shaper, A.G., 1991, Self-assessment of health status and mortality in middle-aged British men, International Journal of Epidemiology 20, 239-245.

Table 1: Definition of the variables used in the analysis

Variable	Definition
Dependent variables	
CATEGORICAL HAPPINESS	Assessment of personal satisfaction (0=the daily life is never a source of personal satisfaction, 1=the daily life is sometimes a source of personal satisfaction, 2=the daily life is a source of personal satisfaction most of the time)
CATEGORICAL HEALTH	Assessment of own health on a three-point scale (0=bad health, 1=fair health, 2= good health)
Independent variables	
MALE	= 1 if male
AGE1	= 1 if age is 18-34 years
AGE2	= 1 if age is 35-44 years
AGE3	= 1 if age is 45-64 years
AGE4	= 1 if age is >64 years
BMI>30	= 1 if body mass index (BMI) is higher than 30[#]
UNEMPL	= 1 if unemployed
SINGLE	= 1 if the individual is not married or cohabiting
HPROBFMS	= 1 if the parents or siblings had any health problems
EDUC1	= 1 if less than high school education
EDUC2	= 1 if high school education
EDUC3	= 1 if university education
INC1	= 1 if the gross annual income is in the first quartile of the income distribution, i.e. <61,500 Skr
INC2	= 1 if the gross annual income is in the second quartile of the income distribution, i.e. 61,500-116,300 Skr
INC3	= 1 if the gross annual income is in the third quartile of the income distribution, i.e. 116,300-161,700 Skr
INC4	= 1 if the gross annual income is in the fourth quartile of the income distribution, i.e. >161,700 Skr
RURAL	= 1 if the individual lives in the countryside or in cities with <30,000 inhabitants
SMALLCITY	= 1 if the individual lives in cities larger than 30,000 inhabitants
BIGCITY	= 1 if the individual lives in Stockholm, Gothenburg or Malmo
HEALTH0	= 1 if the health status is rated as bad health in the categorical health question
HEALTH1	= 1 if the health status is rated as fair health in the categorical health question
HEALTH2	= 1 if the health status is rated as good health in the categorical health question

[#] BMI is defined as: kg/m^2, where kg is the weight in kilograms and m is the height in meters.

TABLE 2: SAMPLE DESCRIPTIVE STATISTICS
(N=NUMBER OF OBSERVATIONS).

VARIABLES	MEAN	SD	MIN	MAX	N
CATEGORICAL HAPPINESS	1.5151	0.6261	0.0000	2.0000	5106
CATEGORICAL HEALTH	1.7393	0.5222	0.0000	2.0000	5106
MALE	0.5027	0.5000	0.0000	1.0000	5106
AGE2	0.3980	0.4895	0.0000	1.0000	5106
AGE3	0.3183	0.4658	0.0000	1.0000	5106
AGE4	0.1443	0.3515	0.0000	1.0000	5106
BMI30	0.0576	0.2330	0.0000	1.0000	5106
SINGLE	0.0447	0.2066	0.0000	1.0000	5106
UNEMPL	0.3263	0.4689	0.0000	1.0000	5106
HPROBFMS	0.1506	0.3577	0.0000	1.0000	5106
EDUC1	0.1645	0.3708	0.0000	1.0000	5106
EDUC2	0.0760	0.2650	0.0000	1.0000	5106
INC2	0.2483	0.4321	0.0000	1.0000	5106
INC3	0.2838	0.4509	0.0000	1.0000	5106
INC4	0.2867	0.4523	0.0000	1.0000	5106
SMALLCITY	0.2244	0.4173	0.0000	1.0000	5106
BIGCITY	0.2777	0.4479	0.0000	1.0000	5106
HEALTH1	0.1808	0.3849	0.0000	1.0000	5106
HEALTH2	0.7793	0.4148	0.0000	1.0000	5106

TABLE 3: ORDERED PROBIT MAXIMUM LIKELIHOOD ESTIMATION RESULTS: DEPENDENT VARIABLE : HAPPINESS (COVARIATES INCLUDED).[a]						
	STRUCTURAL FORM				**REDUCED FORM**	
	HAPPINESS		***HEALTH***		***HAPPINESS***	
VARIABLE	COEFF.	t-VALUES	COEFF.	t-VALUES	COEFF.	t-VALUES
ONE	0.855***	8.532	2.496***	27.845	1.604***	23.685
MALE	-0.115***	-3.172	-0.765E-01*	-1.759	-0.124***	-3.443
AGE2	0.167E-01	0.287	-0.314***	-3.983	-0.374E-01	-0.648
AGE3	0.652E-01	1.022	-0.981***	-12.237	-0.105*	-1.682
AGE4	0.252***	3.740	-1.127***	-13.811	0.303E-01	0.465
BMI>30	-		-0.406***	-5.622	-0.492E-02	-0.073
UNEMPL	-0.220***	-2.659	-0.139E-01	-0.138	-0.218***	-2.670
SINGLE	-0.341***	-8.845	-0.187***	-4.075	-0.365***	-9.572
HPROBFMS	-		-0.103*	-1.913	0.147E-01	0.309
EDUC2	0.120**	2.538	0.339***	5.126	0.165***	3.483
EDUC3	0.242***	3.499	0.471***	5.374	0.307***	4.429
INC2	0.944E-01*	1.800	-0.196E-01	-0.308	0.875E-01*	1.674
INC3	0.852E-01	1.621	0.141**	2.146	0.108**	2.087
INC4	0.139**	2.366	0.411***	5.567	0.206***	3.534
SMALLCITY	0.401E-02	0.092	-0.101E-01	-0.193	-0.101E-04	0.000
BIGCITY	-0.761E-01*	-1.878	-0.119**	-2.494	-0.934E-01**	-2.331
HEALTH1	0.333***	4.171	-		-	
HEALTH2	0.824***	10.917	-		-	
μ_1	1.325***	45.650	1.100***	32.916	1.286***	46.132
N	5106		5106		5106	
Iterations completed	20		21		14	
-Log-L	-4247.116		-2908.374		-4339.038	
Pseudo R^{2b}	0.042		0.099		0.021	
Pseudo R^{2c}	0.070		0.118		0.036	
% corrected	0.587		0.779		0.584	

[a]*** p<.01,** p<.05,* p<.10.

[b]Pseudo R^2 = 1 - LU/LR, where LU is the unrestricted log likelihood values and LR is the restricted log likelihood values (likelihood ratio index).

[c]Pseudo R^2 = 1- exp((-2*(LU-LR/n)), where LU is the unrestricted log likelihood values and LR is the restricted log likelihood values. n is the sample size (Magee 1990).

TABLE 4: PREDICTED PROBABILITIES OF BEING HAPPY MOST OF THE TIME		
VARIABLE	**DIRECT EFFECTS**	**TOTAL EFFECTS**
MALE	0.56369	0.54767
FEMALE	0.58681	0.59647
AGE1	0.56165	0.58923
AGE2	0.56581	0.57463
AGE3	0.57615	0.54802
AGE4	0.61179	0.60095
BMI>30	-	0.57026
BMI=<30	-	0.57219
UNEMPL	0.52905	0.48942
NOT UNEMPL	0.57736	0.57589
SINGLE	0.52766	0.47438
NOT SINGLE	0.59792	0.61820
HPROBFMS	-	0.57696
HPROBFMS=0	-	0.57121
EDUC1	0.56811	0.55220
EDUC2	0.59084	0.61632
EDUC3	0.61192	0.66952
INC1	0.55821	0.52792
INC2	0.57639	0.56257
INC3	0.57458	0.57082
INC4	0.58554	0.60884
RURAL	0.57960	0.58223
SMALLCITY	0.58016	0.58222
BIGCITY	0.56335	0.54546
HEALTH0	0.41651	-
HEALTH1	0.49760	-
HEALTH2	0.60087	-

ADDITIONAL TABLE NOT FOR PUBLICATION

A 1: SURE FEASIBLE GLS ESTIMATION RESULTS: DEPENDENT VARIABLE : HAPPINESS (COVARIATES INCLUDED).[a]

	STRUCTURAL FORM			
	HAPPINESS		***HEALTH***	
VARIABLE	COEFF.	t-VALUES	COEFF.	t-VALUES
ONE	1.127***	20.466	1.942***	73.094
MALE	-0.591E-01***	-3.221	-0.279E-01*	-1.870
AGE2	0.103E-01	0.344	-0.109***	-4.484
AGE3	0.362E-01	1.106	-0.335***	-12.759
AGE4	0.128***	3.674	-0.432***	-15.606
BMI>30	-		-0.187***	-6.298
UNEMPL	-0.122***	-2.931	0.793E-02	0.234
SINGLE	-0.179***	-9.082	-0.683E-01***	-4.270
HPROBFMS	-		-0.373E-01*	-1.937
EDUC2	0.583E-01**	2.420	0.859E-01***	4.383
EDUC3	0.113***	3.360	0.123***	4.511
INC2	0.462E-01*	1.711	-0.122E-01	-0.554
INC3	0.417E-01	1.540	0.556E-01**	2.521
INC4	0.700E-01**	2.339	0.145***	5.985
SMALLCITY	0.141E-02	0.065	-0.122E-01	-0.694
BIGCITY	-0.415E-01**	-2.031	-0.459E-01***	-2.767
HEALTH1	0.203***	4.337	-	
HEALTH2	0.462***	10.449	-	
N	5106		5106	
R^2	0.07500		0.11878	

[a]*** p<.01,** p<.05,* p<.10.