THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 08-09

Roberto Leon Gonzalez
National Graduate Institute for Policy Studies (Japan)
Rimini Centre for Economic Analysis (Italy)

Fu-Min Tseng
University of Nottingham (U.K.)

Socioeconomic Determinants of Mortality in Taiwan: Combining Individual Data and Aggregate Data

Socioeconomic Determinants of Mortality in Taiwan: Combining Individual Data and Aggregate Data

By

Roberto Leon Gonzalez
National Graduate Institute for Policy Studies (Japan)
Rimini Centre for Economic Analysis (Italy)
rlg@grips.ac.jp

Fu-Min Tseng*
University of Nottingham (U.K.)
lexfmt@nottingham.ac.uk

Abstract

There is a very large literature that examines the relationship between health and income. Two main hypotheses have been investigated: the relative income hypothesis and the absolute income hypothesis. Most of previous studies that used mortality data have been criticized for estimating an aggregate model that does not account for non-linear links between health and income at the individual level. In this paper we follow a novel approach to avoid this bias, combining aggregate mortality data with individual level data on socio-economic characteristics. We test the relative and absolute income hypotheses using county level mortality data from Life Statistic of Department of Health and individual level data from Taiwan census FIES for 1976-2003. We find that there is no strong evidence supporting either hypothesis in the case of the general population. In contrast, we find strong evidence that education does have significant effects on individuals' health and the estimates are not sensitive to income equivalent scales.

Keywords: mortality, relative income hypothesis, aggregation bias

* Corresponding author: Fu-Min Tseng, Nottingham School of Economics, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom.

1. Introduction

Many empirical studies have investigated the relationship between income and health and two important hypotheses are the absolute income hypothesis and the relative income hypothesis (for excellent reviews of this wide literature see, for example, Deaton (2003) or Lynch et al. (2004)). The absolute income hypothesis states that health is affected by income, in such a way that other things being constant more income implies better health. On the other hand, the relative income hypothesis states that an individual's health depends especially on her income position in the income distribution. Even though other hypotheses have also been formulated (see for example Wagstaff and Doorslaer (2000)), this paper focuses on these two.

Numerous studies using aggregate data have found support for the relative income hypothesis (e.g. Rodgers (1979), Quick and Wilkinson (1991), Wilkinson (1996), Saunders (1996)). However, they have been criticized by some authors who argued that only individual level data can be used to discriminate between the competing hypotheses (e.g. Gravelle (1998), Wagstaff and Doorslaer (2000), Gravelle, Wildman and Sutton (2002), Wildman, Gravelle and Sutton (2003)). Studies using individual-level data have produced mixed results. While Kennedy, Kawachi and Prothrow-Smith (1998) find evidence that Gini coefficients significantly affect self-rated health, other studies either do not find evidence to support the relative income hypothesis, or find it only in subgroups of the population (e.g. Fiscella and Franks (1997), Daly et al. (1998), Wildman (2002)). However, studies that analyze individual-level data often use self-reported measures of health, which are more prone to measurement error than mortality data. There are studies that analysed individual data on income and mortality (e.g. Fiscella and Franks (1997), Daly et al. (1998), Fiscella and Franks (2000), Lochner et al. (2001), Fritjers, Haisken-DeNew and Shields (2005)). However, individual data on mortality is not available in many countries. In addition, since mortality is a low probability event at the individual level, individual data on mortality provides a limited amount of information, due to the small number of people who die in each wave of the dataset.

In this paper we analyse aggregate mortality data, but following a recent strand of epidemiological literature (Prentice and Sheppard (1995), Salway and Wakefield (2005)), we also use individual level data on income and other socio-economic characteristics to avoid aggregation bias. The econometric approach is similar also to the econometric methods proposed for repeated cross-sections by Deaton (1985) and Browning, Deaton and Irish (1985) (see Cameron and Trivedi (2005, p.p. 770-773) for a summary). The approach consists on first defining a model at the individual level, and then estimating the econometric model that results from aggregating the individual level model over individuals in a county. This approach is made feasible by using individual level data to estimate county averages of regressors. We use county level mortality data from Life Statistic of Department of Health and individual level data from Taiwan census FIES for 1976-2003, and find that there is no strong evidence supporting either income hypothesis.

This paper is organised as follows. Section II discusses the problem of aggregation bias and section III explains the econometric approach followed in this paper to avoid the aggregation bias. Section III also explains the limitations of our econometric approach with respect to using individual-level data for both the dependent and the independent variables. Section IV describes the data and section V the results. Section VI concludes.

2. Discussion of bias

To explain the concept of aggregate bias, consider the following individual level model for the relationship between a measure of health (h_{itk}) and some regressors X_{itk}:

$$h_{itk} = f(\alpha_k + \beta X_{itk}, \varepsilon_{itk}) \qquad (1)$$

where i refers to the individual, t refers to time, k refers to county, ε_{itk} is an unobserved error term and $f(.)$ is a known function. Suppose a researcher does not have observations on individual data for h_{itk} and X_{itk}, but observes instead the county level variables X_{tk} and h_{tk}:

$$X_{tk} = \frac{\sum_{i=1}^{N_k} X_{itk}}{N_{tk}} \qquad h_{tk} = \frac{\sum_{i=1}^{N_k} h_{itk}}{N_{tk}}$$

where N_{tk} is the population size in the k^{th} county in period t. The aggregation bias is likely to arise if the researcher attempts to estimate the following aggregate level model:

$$h_{tk} = f(\alpha_k^* + \beta^* X_{tk}, \varepsilon_{tk}) \qquad (2)$$

Note that the true relationship between X_{tk} and h_{tk}, implied by (1), is:

$$\frac{\sum_{i=1}^{N_k} h_{itk}}{N_{tk}} = \frac{\sum_{i=1}^{N_k} f(\alpha_k + \beta X_{itk}, \varepsilon_{itk})}{N_{tk}} \qquad (3)$$

In general, (2) and (3) are very different. Hence, estimation of (2) will give biased estimates of α_k and β. However, if f was a linear function, then (2) and (3) would be identical, and hence the aggregation bias would not arise.

Gravelle (1998) and Gravelle, Wildman and Sutton (2002) note a typical example in which aggregation bias does arise. Assume that the individual level model is given by:

$$h_{itk} = \alpha_k + \beta_1 I_{itk} + \beta_2 I_{itk}^2 + \varepsilon_{itk} \qquad (4)$$

where I represents income and I^2 is square income. Some cross-country studies might use GDP per capita as a measure of average income, but would not have a measure of the average value of I_{itk}^2. Hence, if (4) defines the true relationship between health and income, a regression of mortality on GDP per capita and GDP per capita squared would suffer from the problem of omitted variable bias (Maddala, 2001, p.p. 159-163), because GDP per capita squared is often not a good approximation of $\sum_{i=1}^{N_{tk}} I_{itk}^2 / N_{tk}$.

3. Econometric Approach

The individual-level model is as follows:

$Y_{itk} = 1$ the individual dies with probability P_{itk}

$Y_{itk} = 0$ the individual survives with probability 1- P_{itk}

where

$$P_{itk} = \beta X_{itk} + \alpha_k \qquad (5)$$

i, t, k denote the i^{th} individual, t^{th} year, and k^{th} county, respectively, and X_{itk} is a vector of regressors. The constant α_k captures the effect of time-invariant county specific characteristics that affect the probability of dying. Note that to simplify the estimation procedure, we make the assumption that the probability of dying depends linearly on the regressors[†], but note that the regressors might include powers of income, age and other variables. Hence, this approach does not rule out the possibility that health and income might be related non-linearly, as advocated by previous studies (e.g. Gravelle, 1998, Gravelle et al. 2002).

The expected value of Y_{itk}, expressed as $E(Y_{itk})$, is equal to 1*Pr(Y_{itk}=1)+0*Pr(Y_{itk}=0)= Pr(Y_{itk}=1). Therefore, the model can be defined as

$$Y_{itk} = \beta X_{itk} + \alpha_k + \varepsilon_{itk} \tag{6}$$

where ε_{itk} is an error with zero mean. Furthermore, the average value of Y_{itk} in county k in year t is

$$Y_{tk} = \frac{1}{N_{tk}} \sum_{i=1}^{N_{tk}} (Y_{itk})$$

where N_{tk} is the population size in the k^{th} county and t^{th} period. Note that Y_{tk} is the mortality rate in county k in period t. Hence, (6) implies that Y_{itk} can be expressed as:

$$Y_{tk} = \beta X_{tk} + \alpha_k + \varepsilon_{tk} \tag{7}$$

where X_{tk} is a vector containing the average values of the regressors in X_{itk}:

$$X_{tk} = \frac{1}{N_{tk}} \sum_{i=1}^{N_{tk}} (X_{itk})$$

and:

$$\varepsilon_{tk} = \frac{\sum_{i=1}^{N_{tk}} \varepsilon_{itk}}{N_{tk}}$$

In this study, we have data on Y_{tk} (mortality data) and we will use sample averages of X_{itk}, $\hat{X}_{tk} = \frac{1}{\tilde{N}_{tk}} \sum_{i=1}^{\tilde{N}_{tk}} (X_{itk})$, with $\tilde{N}_{tk} < N_{tk}$, to proxy for X_{tk}. Given the large sample sizes that we use to calculate these averages (see Table 1), the bias introduced because of measurement error is negligible (Prentice and Sheppard (1995), and Salway and Wakefield (2008)). Prentice and Sheppard (1995) by means of a Monte Carlo experiment found that when $\tilde{N}_{tk} = 100$, the bias was about -2.0%, with a 95% confidence interval of (-3.8%, -0.1%). Moreover, they found that the correction they proposed to reduce the

[†] Non-linear links between the probability and the set of regressors could also be considered following the approach in Prentice and Sheppard (1995).

measurement bias, which is superior asymptotically, in practice was not worthy when $\tilde{N}_{tk}$ was at least 100. The Monte Carlo experiment carried out by Salway and Wakefield confirms that the measurement error bias would be negligible in our case. They found that even when $\tilde{N}_{tk}$ is as small as 25, the bias is only about 6%.

Note that the main difference in practice of our approach with other aggregate data studies is that the latter implicitly assume that $E(X^2)$ can be approximated with $(E(X))^2$, where X denotes an explanatory variable, and E(.) is the expected value operator. We do not make this assumption, while the aggregate level model (7) is still consistent with the individual level model (5), which allows for non-linear links between health and other variables.

We use fixed effects and random effects models to estimate the parameters and compare the results. However, it is possible that the error term ε_{tk} has heteroskedasticity and/or autocorrelation. Hence, we use robust fixed effects estimation to correct this bias. As for random effects models, we use robust population averaged estimation to account for heteroskedasticity and autocorrelation.

Scope and Limitations

The aggregate approach proposed in Prentice and Sheppard (1995) produces consistent estimates of the parameters of the individual-level model under a set of assumptions which is more restrictive than would be necessary if individual-level data were available. The crucial extra assumption is that the error term in the individual-level model is uncorrelated with the regressors. This assumption could be relaxed to some extent by using fixed effects and individual-level data. That is, individual-level data would allow us to effectively control for all individual characteristics that are time-invariant, even for those that are not observed. In the aggregate approach described here, however, we can only control for observed individual characteristics (e.g. income, age, education, occupation, gender, etcetera).

Another assumption is that $\hat{X}_{tk}$ is a consistent estimator of X_{tk} and that $\tilde{N}_{tk}$ is sufficiently large. This requires that the sample used to calculate $\hat{X}_{tk}$ is representative of the population in that county and year. Since this sample is taken at a particular date within the year, we need to assume that $\hat{X}_{tk}$ is roughly constant during the year. Note that in our case we will be using income adjusted by the number of household members as one of the regressors. Thus, if X_{tk} represents the yearly average, we implicitly assume that the probabilities of death and birth do not vary over the year as much as to make the estimator $\hat{X}_{tk}$ inconsistent.

Year	Minimum	5%	Mean	95%	Maximum
1	291	896	2357	5637	6755
2	286	850	2342	5211	6740
3	453	1110	3344	8155	9590
4	457	1077	3269	8474	9555
5	458	1089	3392	9165	10037
6	449	1113	3490	9902	9957
7	444	981	3249	10050	10465
8	449	989	3364	10976	10989
9	407	969	3349	11011	11042
10	400	895	3283	10795	11143
11	359	941	3236	10695	11139
12	381	871	3155	8778	10718
13	331	840	3061	8576	10091
14	338	828	3038	8420	10155
15	294	752	2993	8509	9853
16	289	830	2975	8358	10001
17	288	760	2930	8386	9845
18	235	728	2922	8438	9959
19	258	682	2868	8153	9929
20	399	788	2508	7459	9332
21	425	711	2328	6891	9470
22	374	817	2282	5771	9468
23	371	843	2287	5782	9036
24	379	873	2170	5582	8981
25	367	815	2164	5278	9145
26	353	805	2104	5477	7169
27	595	839	2152	5504	6889
28	617	743	2085	5302	6744
29	628	800	2061	5323	6437

Table 1: Number of values to calculate county averages of regressors: Minimum value, 5% percentile, Mean, 95% percentile and Maximum value.

4. Data

The dependent variable is the population mortality rate and the regressors at the individual level are: age, age squared, a dummy for gender, 4 dummies for educational achievement, 6 dummies for occupation, log of equivalised disposable income, square of log equivalised income and the Gini coefficient[‡]. Income refers to disposable household income divided by two income equivalent scales. One is the number of members in the household, named income scale 1 and the other one is as the formula: $(adult + 0.5 \times children)^{0.9}$, named income equivalent 2. In order to estimate equation (7), individual level data is used to estimate the area averages of these regressors.

There are two data sources that we use in this paper. One is the data on mortality rate obtained from Life Statistic of Department of Health of Taiwan. From this dataset, we select crude mortality rate (CMR), which is the total number of deaths per 1000 people in a year.

The number of total districts in Taiwan now is twenty three. These twenty-three districts include sixteen counties, five cities, and two municipalities governed directly under the jurisdiction of the Central Government (Taipei city and Kaohsiung city). However, our panel data is unbalanced because there were only twenty-one counties before 1982.[§]

The other data source we use is Taiwan Family Income and Expenditure Survey. We select 29 years data from 1976 to 2004. We calculate averages, which are needed to estimate equation (7), using all available households in the survey.

The number of household samples drawn from each district was proportional to its population size. The number of households in the survey varies from year to year. The smallest total sample size, in terms of households, is 9033 in 1977 and the largest one is 16435 in every year from 1983 to 1994. The Family Income Expenditure Survey contains data on the members of each household. The smallest total sample size, in terms of number of individuals, is 47411 in 2004 and the largest one is 77393 in 1983.

[‡] A model using log income instead of income was also estimated and results were qualitatively the same.
[§] These two extra districts are Hsinchu city and Chiayi city, which were two towns that belonged to Hsinchu county and Chiayi county originally. Because the population in these two cities grew, they were upgraded to the same level as county. Roughly, Hsinchu county and Chiayi county maintain the same scale of territory.

The explanatory variables include demographical variables and income-expenditure variables. The demographical variables comprise gender, age, age squared, education, and occupation. The income variables include log of disposable income, square of log disposable income and the Gini coefficient.

The gender variable in the county-level equation (7) is the proportion of males in one particular district and year. Similarly, the variables of education and occupation in the county-level equation (7) are proportions. We categorize education into 5 groups according to the number of years enrolled at school and occupation into 7 groups according to the type of job. The five groups in education are less than 1, 1-6, 7-9, 10-12, and more than 12. Note that these variables enter the individual level equation (5) as dummy variables, which imply that they enter equation (7) as proportions. Because the sum of these proportions equals one, only four of them are included in the model. With respect to occupation, the seven groups are: 1. Professionals, 2. Clerks, 3. Technicians and associate professionals, 4. Service workers, shop and market sales workers, 5. Agricultural, animal, husbandry, forestry, and fishing workers, 6. Product machine operators and related workers, 7. Unemployed.

These occupation variables are denoted from occu1 to occu7. Of course, the sum of these seven variables equals to one too, so only 6 of them enter the regression. In order to measure income, we use household disposable income over the number of household members. The definition of household disposable income is total receipts minus non-consumption expenditure.[**] We then calculate the average of household disposable income and household disposable income squared as explained near to equation (7). As to the Gini coefficient, it is computed using the individual disposable income in every year. Table 1 shows the average Gini coefficient in each district over 29 years.

Finally, the number of observations that we use to estimate the panel regression is 655 (23 districts and 29 years minus 12 observations which are missing because two counties have no data from 1976 to 1981).

[**] Total receipts include six terms. They are compensation of employees, entrepreneurial income, property income, imputed rent income, current transfer receipts, and miscellaneous receipts. The non-consumption is composed of interest and current transfer expenditures.

Table 1. The mean of Gini coefficient in each district

District	Mean [Std. Dev.]	District	Mean [Std. Dev.]
Taipei County	0.275 [0.017]	Pingtung County	0.274 [0.02]
Yilan County	0.278 [0.024]	Taitung County	0.306 [0.033]
Taoyuan County	0.267 [0.013]	Hualien County	0.317 [0.022]
Hsinchu County	0.265 [0.015]	Penghu County	0.304 [0.032]
Miaoli County	0.267 [0.05]	Keelung City	0.271 [0.022]
Taichung County	0.269 [0.02]	Taichung City	0.287 [0.021]
Changhua County	0.28 [0.014]	Tainan City	0.277 [0.015]
Natou County	0.298 [0.019]	Kaohsiung City	0.278 [0.012]
Yunlin County	0.276 [0.024]	Taipei City	0.294 [0.035]
Chiayi County	0.289 [0.017]	Hsinchu City	0.321 [0.027]
Tainan County	0.28 [0.02]	Chiayi City	0.307 [0.02]
Kaohsiung County	0.273 [0.014]		

† The mean of each district is the average Gini coefficient for 29 years except for Hsinchu City and Chiayi City because these two city are only with 23 years.

5. Result

We use robust fixed effect estimation (Wooldridge, 2002, Section 10.5.4), which allows for heteroskedasticity and (intra-group) autocorrelation in the error term. We also present a random-effects type estimation, using a robust population averaged method (Wooldridge, 2002, Section 10.4.2) that assumes, unlike the fixed effects, that the time-invariant unobserved variables are uncorrelated with the explanatory variables. However, as the robust fixed effects estimation, it allows for heteroskedasticity and autocorrelation in the error term. Results are shown in Table 2. The results with random and fixed effects are similar except for the occupational variable.

None of the income variables have a significant effect on the probability of dying. Thus, we find no evidence for either the absolute or relative income hypotheses. However, education variable, Edu5, is significant at a 5% level. Individuals with more than 12 years of education have lower probability of dying (holding other things constant). In particular, the probability of dying for an individual with more than 12 years of education is between 0.56% and 0.59% smaller than the probability for an individual whose education level is less than one year. However, in the random effect model, individuals who are professionals have a higher probability of dying and the probability of dying is 0.68% higher for professionals compared to individuals who are unemployed when the other variables are held constant.

Table 2. Robust estimation with income equivalent scale 1†.

Dependent Variable	Fixed Effects Coef.	Fixed Effects Robust Std. Err.	Population Averaged Coef.	Population Averaged Semi-robust Std. Err.
Mean of log disposable income	-1.591	1.276	-1.639	1.242
Mean of square log disposable income	0.089	0.064	0.092	0.063
Gini coefficient	0.246	0.213	0.193	0.957
Age	-0.394	0.213	-0.393	0.204
Square of age	0.007*	0.003	0.007**	0.003
Gender	1.617	2.498	1.627	2.482
Education				
Edu 2(1 - 6 years)	-1.646	1.278	-1.552	1.187
Edu 3(7 – 9 years)	1.049	2.397	1.336	2.226
Edu 4(10 – 12 years)	-0.346	1.969	-0.104	1.878
Edu 5(More than 12 years)	-5.729**	1.806	-6.044**	1.761
Occupation††				
Occu1	6.512	3.295	6.706*	3.149
Occu2	-0.915	2.587	-1.351	2.527
Occu3	0.145	3.144	-0.416	3.05
Occu4	-1.616	1.771	-1.887	1.725
Occu5	-1.03	1.339	-0.586	1.292
Occu6	2.378	2.048	1.67	1.871
Constant	14.841	7.387	14.896*	7.229

† The equivalent scale is the number of household members.
†† Occu1: professionals; Occu2: clerks; Occu3: technicians and associate professionals; Occu4: service workers, shop and market sales workers; Occu5: agricultural, animal, husbandry, forestry, and fishing workers; Occu6: product machine operators and related workers. The comparative group is unemployment.
††† * significant at 5% level; ** significant at 1% level

As expected, age appears as a significant determinant of the probability of dying. The relationship is nonlinear and the turning point of age is 28. An increase of age induces a decrease in the individual probability of dying before the age of 28. After this age, the relationship becomes positive. However, there is no significant evidence for the effect of gender. Table 3. shows that results are similar when the income equivalent scale 1 is replaced by income equivalent scale 2.

Our results on education variables are coherent with the intuition that people with higher education level have more knowledge and ability to look after themselves or others. A large body of literature also supports that the educational attendance has a positive association with health.

Table 3. Robust estimation with income equivalent scale 2†.

Dependent Variable	Fixed Effects		Population Averaged	
	Coef.	Robust Std. Err.	Coef.	Semi-robust Std. Err.
Mean of log disposable income	-1.251	1.389	-1.291	1.354
Mean of square log disposable income	0.071	0.067	0.073	0.066
Gini coefficient	0.036	1.055	-0.007	0.991
Age	-0.394	0.214	-0.393	0.204
Square of age	0.007*	0.003	0.007**	0.003
Gender	1.703	2.485	1.714	2.47
Education				
Edu 2(1 - 6 years)	-1.627	1.313	-1.534	1.213
Edu 3(7 – 9 years)	1.029	2.395	1.309	2.227
Edu 4(10 – 12 years)	-0.31	2.018	-0.066	1.927
Edu 5(More than 12 years)	-5.562**	1.824	-5.874**	1.778
Occupation††				
Occu1	6.574	3.319	6.767*	3.171
Occu2	-0.525	2.565	-0.948	2.513
Occu3	0.195	3.126	-0.366	3.032
Occu4	-1.518	1.754	-1.789	1.711
Occu5	-0.991	1.335	-0.546	1.291
Occu6	2.312	2.065	1.607	1.883
Constant	13.081	8.119	13.087	7.969

† The formula of equivalent scale is (number of adult + 0.5* number of children)^0.9.
†† Occu1: professionals; Occu2: clerks; Occu3: technicians and associate professionals; Occu4: service workers, shop and market sales workers; Occu5: agricultural, animal, husbandry, forestry, and fishing workers; Occu6: product machine operators and related workers. The comparative group is unemployment.
††† * significant at 5% level; ** significant at 1% level

6. Conclusion

In this paper we analysed the socio-economic determinants of mortality, with a particular focus on the absolute and relative income hypotheses, using a novel approach to avoid aggregation bias. Following a recent strand of epidemiological literature (Sheppard and Prentice (1995), Salway and Wakefield (2005)), we combined individual level data on income and other socio-economic characteristics with aggregate data on mortality. When compared with using individual-level data for both mortality and regressors, the proposed approach has the disadvantage that it cannot control for unobserved time-invariant characteristics. However, it has an advantage over the aggregate studies that have neglected non-linear links at the individual-level data model. We analysed Taiwanese data and found no evidence to support either the absolute or the relative income hypotheses. However, results confirm the positive effects of education on the health of individuals although the evidence on occupation related effects on health is mixed. In addition, we also use different income equivalent scales and find that it is not a sensitive factor in our analysis.

References:

Browning, M., Deaton, A. and Irish, M. (1985), A profitable Approach to Labor Supply and Commodity Demands over the Life-Cycle. Econometrica, 53, 503-543.

Cameron, A.C. and Trivedi, P.K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press, Cambridge.

Daly, M. C., Duncan, G.J. , Kaplan, G. A. and Lynch, J. W. (1998), Macro-to-Micro Links in the Relation Between Income Inequality and Mortality. *Milbank Quarterly* 76, 315-39.

Deaton, A. (1985). Panel Data from Time Series of Cross-Sections. *Journal of Econometrics*, 30, 109-126.

Deaton, A. (2003). Health, Inequality, and Economic Development. *Journal of Economic Literature*, 41, 113-158

Fiscella, K., Franks, P., (1997), Poverty or income inequality as predictors of mortality. *British Medical Journal* 314, 1724–1727.

Fiscella, K., Franks, P., (2000), Individual Income, Income Inequality, Health, and Mortality: What Are the Relationships? *Health Services Research*, 35, 308 - 318.

Frijters, P., Haisken-DeNew, J., and Shields, M.A., (2005). Socio-Economic Status, Health Shocks, Life Satisfaction and Mortality: Evidence from an Increasing Mixed Proportional Hazard Model. *IZA Discussion Paper*, 1488.

Gravelle, H. (1998). How much of the relationship between mortality and unequal distribution of income is a statistical artefact? *British Medical Journal*, 314, 382-385.

Gravelle, H., Wildman, J. and Sutton, M. (2002). Income, income inequality and health: what can we learn from aggregate data? *Social Science & Medicine*, 54, 577- 589.

Judge, K. (1995). Income distribution and life expectancy: a critical appraisal. *British Medical Journal*, 311, 1828-1285.

Judge, K., Mulligan, J.A. and Benzeval, M. (1998). Income inequality and population health. *Social Science & Medicine, 46, 567-579.*

Leigh, A. and Jencks, C. (2007). Inequality and mortality: long-run evidence from a panel of countries. *Journal of Health Economics*, 26, 1-24.

Lochner, K., E. Pamuk, D. Makuc, B. P. Kennedy, and I. Kawachi. (2001) "State-Level Income Inequality and Individual Mortality Risk: A Prospective, Multi-level Study." *American Journal of Epidemiology* 91, 385-391.

Lynch, J., Smith, G.D., Harper, S. and Hillemeier, M. (2004.) 'Is Income Inequality a Determinant of Population Health? Part 1. A Systematic Review'. Milbank Quarterly, 82(1): 5-99

Maddala, G.S., 2001. *Introduction to Econometrics*. Third Edition, Wiley.

Prentice, R. and Sheppard, L. (1995). Aggregate data studies of disease risk factors. *Biometrika*, 82, 1, 113-25

Quick, A., and Wilkinson, R. G. (1991). Income and health. *Socialist Health Association*.

Rodgers, G. B. (1979). Income and inequality as determinants of mortality: an international cross-section analysis. *Population Studies*, 33, 343-351.

Salway, R. and Wakefield, J. (2008). A hybrid model for reducing ecological bias. *Biostatistics,* 9(1), 1-17.

Saunders, P. (1996). Poverty, income distribution and health: an Australian study. No.128, Reports and Proceedings. Social Policy Research Centre, University of New South Wales.

Sheppard, L. and Prentice, R. (1995). On the reliability and precision of within- and between- population estimates of relative rate parameters. *Biometrics*, 51, 853-863.

Wagstaff, A. and Doorslaer, E. V. (2000). Income inequality and health: what does the literature tell us? *Annu. Rev. Public Health*, 21, 543-67.

Wildman, J., (2003). Income related inequalities in mental health in Great Britain: analyzing the causes of health inequality over time. *Journal of Health Economics*, 22, 295-312

Wildman, J. Gravelle, H. and Sutton, M. (2003), Health and income inequality: attempting to avoid the aggregation problem, *Applied Economics*, 35, 999-1004

Wilkinson, R.G., (1996) *Unhealthy Societies: The Afflictions of Inequality*. Routledge & Kegan Paul, London.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, Massachusetts.