The University of Manchester

MANCHESTER 1824

Economics Discussion Paper Series EDP-1110

Inequality of Happiness in US: 1972-2008

Indranil Dutta
James Foster

February 2011

Economics
School of Social Sciences
The University of Manchester
Manchester M13 9PL

Inequality of Happiness in US: 1972-2008[1]

Indranil Dutta[2]
Department of Economics, University of Manchester, UK.

James Foster
Department of Economics, George Washington University, USA.

ABSTRACT

It is well accepted that a country's GDP may not reflect its level of well-being. In recent years, happiness has emerged as an alternative indicator of well-being and research so far has focussed on determining the level of happiness. While it is important to look at the level, the distribution of happiness is also a salient aspect in any evaluation of well-being. This has so far been overlooked in the literature. Our paper is an attempt to fill this gap, by measuring the inequality of happiness for US from 1972 to 2008 using the General Social Survey database. The data on happiness, however, is ordinal and any attempt in using existing 'mean' based measures of inequality will be problematic. Based on the methodology developed in Allison and Foster (2004) we are able to unambiguously rank the happiness distribution over the years. It also allows us to overcome the issue of 'ordinality' in the data through a median centered approach. Further we decompose the median inequality measure of happiness across gender, race and region.

Key Words: Dominance, Happiness, Inequality, Median Inequality, Scale Independence.
JEL Classification Number: D39, I39

February, 2011

[1] We are grateful to Angus Deaton for suggesting us this research topic.
[2].Corresponding author. Address: School of Social Sciences, Arthur Lewis Building, Oxford Road, Manchester, M13 9PL. Email: i.dutta@manchester.ac.uk. Telephone: +44 (161) 275 4860

1. Introduction

It is well accepted that a country's GDP per capita does not reflect its level of well-being (Sen 1987). While well-being is a broad concept, difficult to pin down, it is increasingly being measured in terms of subjective well-being indicators such as happiness (Oswald 1997, Ng 1997). Research so far has focussed on finding levels of happiness in different countries over time and the factors that drive it (Alesina et al 2004, Di Tella et al 2003). Although understanding the level of happiness is important, it is equally important, if not more, to understand the distribution of happiness, especially if we take happiness to be a broader indicator of well-being. In other words, the inequality of happiness is a salient aspect in any evaluation of well-being, which, however, has not received much attention in the literature. Our paper is an attempt to fill this gap, by measuring the inequality of happiness for US.

Interest in happiness is to some extent spurned by rich data sets spanning over last three decades. When it comes to the subjective evaluation of wellbeing the current surveys typically include questions which asks respondents to classify their current level of happiness in one of three or four categories of happiness ranging from `very happy' to `not at all happy'. Therefore, the data on happiness is in an ordinal scale, i.e. we can tell that `very happy' is better than `not at all happy' but by how much it is better is not known. Since we are especially interested in the `spread' of the distribution, the information on the difference between the two categories becomes crucial. In absence of this information, use of most of the standard measures of inequality will be problematic.

Kalmijin and Veenhoven (2006) measure the inequality of happiness across nations. They assume a cardinal scale across the categories of happiness and thus are able to use and evaluate the standard measures of inequality such as standard deviation and the Gini coefficient. They recommend using the standard deviation, inter-quartile range and the absolute Gini index and the mean absolute difference for measuring inequality of happiness. We demonstrate later that none of these indices are suitable to measure inequality under ordinal scale. But more importantly, one can question the appropriateness of using a single cardinal scale when dealing with ordinal data as they have done.

One justification for using cardinal scale has been put forth in van Praag (1991) who argued that when choosing between different categories, respondents do have a numerical scale in mind although the numerical scale may differ between respondents. In other words, we can associate an ordinal scale with at least an interval scale. One major issue with the study is that all respondents were provided a common numerical scale and were thus able to associate the given ordinal categories to some range in the numerical scale. However, it is not clear in the absence of a common numerical scale, how respondents will relate the ordinal categories to a numerical scale. But even if we agree that behind the categorical answers is a numerical scale, surveys do not report any numerical scale associated with the categories. In the absence of such information, it is, thus, not very satisfactory to use a cardinal scale.

In a recent paper Stevenson and Wolfers (2008) calculate the inequality of happiness for US based on General Social Survey (GSS). They, however, cardinalize the ordinal data by assuming that happiness is derived from an implicit Probit distribution. The variance of that distribution, derived based on the percentage of people in the different happiness categories, is referred to as the inequality of happiness. Apart from using a particular distribution to cardinalize the variable, they also assume that the distribution itself remains the same over all the years. But more problematically variance as a measure of inequality is unsuitable particularly in this context.

To overcome the problems of measuring inequality under an ordinal scale we employ the methodology developed by Allison and Foster (2004) which they used to measure inequality in self reported health status. Similar to happiness, self reported health also had an ordinal scale. We use this new methodology for the first time in the context of measuring inequality of happiness. This not only allows us to measure inequality of happiness more satisfactorily but in the context of happiness in the US it allows to explore some interesting methodological extensions.

The plan of the paper is as follows: the next section discusses the problems with the standard measure and explains the Allison-Foster methodology in the context of happiness; in Section 3, we apply this new methodology to the happiness data from the General Social Surveys (GSS) in US and discuss the ranking of the different years in terms of lower happiness inequality; Section 4 uses decomposes the Allison-Foster measure of inequality across gender, race and regions. The last section concludes the paper.

2. Measures of Inequality with Ordinal Data

2.1 Problems with Standard Inequality Measures

When it comes to ordinal or categorical data, standard measures of inequality such as the standard deviation, coefficient of variation and Gini coefficient may not be reliable. This is because in order to derive the level of inequality using these measures, the categorical data needs to be scaled. For instance when it comes to the level of happiness, in the US, people have three categories to choose from: Very Happy, Pretty Happy and Not at all Happy. Although it is obvious that Very Happy is better than Pretty Happy which in turn is better than Not at all Happy, to measure inequality we need to know how far apart these categories are from each other. We may use a scale of (1,2,3), where Very Happy is at 3, Pretty Happy at 2 and Not too Happy at 1. Once the scale is established, the standard measures can be applied to derive the level of inequality. There is, however, no reason that we should be restricted to one particular scale. We can take another scale, for example, of (1,2,5), where Very Happy is associated with 5, Pretty Happy with 2 and Not too Happy with 1. The levels of inequality calculated using the standard measures will clearly change, what is worrying is that the orderings of the levels of inequality between groups or across time, or across geographical boundaries will change too. Hence, for instance for US, while 1990 may have been a year of less happiness inequality compared to say 2000 under one scale, yet another scale may show the opposite. In such circumstance, it becomes very difficult to understand whether inequality is increasing or decreasing over time, across regions and between groups.

For most of the inequality measures the source of the problems comes from the fact that the deviations from the mean will not be order preserving since the mean itself is not order preserving under scale changes (Allison and Foster 2004). One may, however, point out that measures such as the absolute Gini or the inter-quartile range will be free from this criticism since they are not mean dependent. As it turns out, for the absolute-Gini, which is the sum of all the pair wise differences, if we change the scales associated with each category by the same amount, the inequality ordering will remain unchanged Therefore, if we change the scale from (1,2,3) to (3,4,5), the absolute-Gini will be order preserving. This is because under such transformation of the scale the difference between each pair remains unchanged. On the other hand if the scales are changed to (3,4,7) instead, then there is no guarantee that the absolute-Gini will continue to preserve the ranks. Since there is no reason why we should be considering only one kind of scale transformation, the use of absolute-Gini becomes problematic.

The inter-quartile range on the other hand looks at the difference between the first and the third quartile (or in other words the difference between the 75th percentile and the 25th percentile). This is clearly order preserving for any transformation of the scales. But it is subject to the criticism that, like the measure of range, this does not capture the full extent of inequality. Take, for instance, the following distribution of happiness: (1,3,0), i.e. in the context of our example, out of four people one report Not too Happy and the rest report Pretty Happy. Let the scale be c=(1,2,3). Now suppose with the scales remaining unchanged, the distribution changes to: (1,2,1), i.e. one of the person who was reporting Pretty Happy now reports Very Happy. Clearly the inequality has increased, yet the inter-quartile range for both these distributions is the same i.e. 1.

Another measure of mean independent inequality would be `the percentage outside modus'(Kalmijin and Veenhoven 2006). In other words, inequality is 1 minus the share of the population in the modal category. This measure, however, suffers from the same flaw as the inter-quartile range in that it is not always sensitive to increase in inequality. Consider the following distribution (1,3,1) where one person is reporting Not too Happy, three are reporting Pretty Happy and one is reporting Very Happy with a linear scale c=(1,2,3). The modal category is Pretty Happy. The proportion of population outside the modus, and thus the level of inequality, is 0.4. Now consider another distribution (0,3,2) where three are reporting Pretty Happy and two are reporting Very Happy. Compared to the previous distribution, it indicates a lowering of inequality in the sense that the dispersion is lower in this distribution. Yet there is no change in inequality since the proportion of people outside the modus remains the same. There are also other issues that arise with mode based measures of inequality: there may be distributions where there is no mode, or there may be multiple modes. In such cases it is not very clear how the inequality should be measured.

Given these issues with both mean based and mode based measures of inequality, Allison and Foster (2004) propose a median based dominance concept to evaluate inequality on categorical data. The next section describes their methodology.

2.2 The Allison-Foster dominance criteria

Allison and Foster (2004) note that one of the measures of central tendency that is order preserving under scale transformations is the median. The median, therefore, becomes a natural choice from which to evaluate the dispersion in the distribution. Although the Allison-Foster methodology is quite general and can be applied in wider contexts involving categorical data, we illustrate their measure in the context of happiness, where there are only three categories of Very Happy, Pretty Happy and Not at all Happy.

Consider any two happiness distribution, X and Y, both with the same median category. If the cumulative distribution of X in each of the three categories is lower than the cumulative distribution of Y, then invariably X is a better distribution than Y, or in the other words X would first order dominate Y (i.e. X F-dominates Y). The intuition is that X has lower percentage of people in the inferior categories and more percentage in the better categories.

Allison and Foster (2004) propose the S-dominance concept to evaluate inequality in an ordinal setting. We illustrate briefly the concept of S-dominance with the help of Figure 1 below.

[Insert Figure 1]

The broken lines represent the cumulative distribution of Y and the solid line represents the cumulative distribution of X. Suppose the common median category is Pretty Happy. In Figure 1, distribution X has a greater population share in the category below the median and a greater population share in the category above the median compared to Y. Since X has a greater mass in both the tails of the distribution compared to Y, X will have higher inequality, i.e. Y is S-dominant over X. This is because, in an intuitive sense it means that more people are concentrated in the middle for distribution X than distribution Y, or in other words the `spread' of the distribution is lower for X. In such circumstance it is quite possible that although X has lower inequality than Y, it also has lower `average' achievements in terms of happiness than Y.

Let f^1_X and f^3_X represent the share of the population in category 1 (Not too Happy) and category 3 (Very Happy) respectively for distribution X. Similarly for distribution Y, f^1_Y and f^3_Y represent the share of the population in category 1 and 3 respectively. If $f^1_X>f^1_Y$ and $f^3_X>f^3_Y$ then X has more mass in the tails and hence will have more inequality. On the other hand if $f^1_X<f^1_Y$ and $f^3_X<f^3_Y$, then Y has more mass in the tails and will have more inequality compared to X. When the inequalities are in conflict then we may not be able to rank the distributions according the S-dominance. So when $f^1_X>f^1_Y$ and $f^3_X<f^3_Y$, S-dominance cannot rank the distributions, however, in this case Y first order dominates X. Similarly if $f^1_X<f^1_Y$ and $f^3_X>f^3_Y$, then X first order dominates Y, although the distributions cannot be compared using S-dominance. These can be easily summarised in the table below

[Insert Table 1]

Therefore, for the three category case, either a S-dominance relationship or a F-dominance relationship will hold. It implies that S-dominance and F-dominance taken together will completely rank and set of distributions with three categories case, however, such may not be true of distributions with more than three categories.

Note that these two dominance concepts of First order dominance and S-dominance are clearly scale independent since they are based on the frequency distribution. So what ever the scale may be, so long as the distributions do not change, the ordering will always be preserved.

2.3 Median Based Measures of Inequality

Allison and Foster (2004) also propose a measure of inequality which is the difference between the `mean' of the upper half of the distribution and the `mean' of the lower half of the distribution. The mean, however, is dependent on the scale used. It is best to illustrate the concept through an example. Consider a distribution X:(1,3,1) where one individual reports Not too happy, three reports Pretty Happy and one reports Very Happy. If the scale used is (-1,0,1) where -1 stands for Not too happy, 0 for Pretty Happy and 1 for Very Happy, then the mean in the lower half of the distribution would be -0.2. The mean of the upper half of the distribution would be 0.2. The overall inequality under the proposed measure would be 0.4.

To represent the measure in more general terms suppose there are n observations with N categories and mth category is the median category. Let f^i_J be the percentage of people in the ith category and F^i_J be the cumulative percentage of people in the ith category of distribution J. Consider any scale $c=(c_1,c_2,...,c_N)$. Then the Allison-Foster measure of inequality, can be written as:

$$I_{AF} = (\sum_{i>k} c_i.f^i + c_m.(F^m-0.5)/0.5) - (\sum_{i<m} c_i.f^i + c_m.(0.5-F^m)/0.5).$$

The first and the second term in the above equation is the mean of the upper and lower half of the distribution respectively. The difference yields the Allison-Foster measure. The above equation can be further reduce to

$$I_{AF}=2(\sum_{i>m} c_i.f^i - \sum_{i<m} c_i.f^i + c_m.(F^m + F^{m-1} - 1)).$$

For the three category case, with the second category being the median category, the above will be reduced to

$$I_{AF}=2(c_3.f^3 - c_1.f^1 + c_2.(F^2+F^1-1)).$$

Clearly the level of inequality and the orderings would be very much dependent on the scale. Allison and Foster (2004) prove that if for every scale c, the inequality based on mean difference indicates that distribution X has higher inequality than Y, then it is indeed the case that Y S-dominates X. Hence, if for a particular scale X is shown to have a higher inequality than Y, then there is a possibility that Y indeed S-dominates X. It, however, has to be noted that since the proposed measure is not scale independent, there is a possibility that different scales may lead to conflicting orderings.

Abul Naga and Yalcin (2008) expand on the Allison-Foster inequality measure by considering a weighted difference between the percentage of people in the lower half of the distribution and the upper half of the distribution. If inequality depends on the population of the upper and the lower half of the distribution then how much importance should be given to the lower and upper half in the over inequality measure is a subjective judgement. This generalisation thus builds the subjective judgements in to the inequality measure and can be expressed as

$$I_{AN}(\alpha,\beta) = f(\sum_{i>m}(F^i)^\alpha, \sum_{i\geq m}(F^i)^\beta),$$

where α, β reflect the value judgements of the society.

For a three category case and when $\alpha=\beta=1$, the inequality measure is

$$I_{AN}(1,1) = 1.5-((\sum|F^i - 0.5|).$$

If the second category is also the median category then,

$$I_{AN}(1,1) = 1-f^2.$$

Effectivelywhat it means is that percentage of people not in the median category would be considered as the level of inequality.

In the next section we use these concepts of dominance and the proposed inequality measure to calibrate the level and the inequality of happiness in the US.

3. Empirics

3.1 Data

We use the General Social Surveys in US, which collects extensive data on individual levels of happiness along with a rich set of personal information. The data is for 27 years from 1972 to 2008, with a gap of few years in between. In each of these years a nationally representative sample was chosen. Our main interest is in the following GSS question: "Taken all together, how would you say things are these days -- would you say that you are very happy, pretty happy or not too happy?" The question is asked to the head of the household. The average age of the respondents was 45 years. It does not contain information about the happiness levels of other adults in the household. It should also be noted that this is not a panel survey which means that it does not track the same group of people over the different years.

The total number of available responses to the happiness questions is 48,318. There were, however, over sampling of the black population for the years 1982 and 1987. Following Stevenson and Wolfers (2008) we have dropped the over sampled observations and the interviews that took place in Spanish in 2006. This reduces the total responses to 47,394. We weighted the observations according to the sample weights (WTSALL) which were provided. The weighted distribution of the responses over the three categories of happiness is summarised in the Table 2 below.

[Insert Table 2]

The middle category of Pretty Happy turns out to be the median category. We have checked for each year separately and in each case Pretty Happy is the median category. This ensures that the data satisfies the condition of the Allison-Foster methodology and hence their techniques can be applied.

However Stevenson and Wolfers (2008) found that question order effects can have an impact on how the respondents classify themselves within the happiness categories. Out of the 27 years, this was an issue for 5 years: 1972, 1980, 1985, 1986, 1987. They corrected for it and presented the revised percentages in each of these categories for these years. In the next sections we shall present the results first for the weighted data and then for the question order corrected weighted data.

3.1 Inequality of Happiness

As noted before, when there are three categories, just knowing the share of the population in the lowest and the highest category will allow us to rank the different distributions. It will also allow us to calculate the mean of the lower tail and upper tail of the distribution, for any given scale. The difference of the means would lead us to the Allison-Foster measure of inequality. We present the calibrations for the weighted US data from 1972 to 2008 in Table 3a below.

[Insert Table 3a]

The first three columns shows the share of the population in the lowest category (Not too Happy), the median category (Pretty Happy) and the highest category (Very Happy) respectively. The next two columns contains the mean of the upper half and the lower half of the distribution respectively. The fifth and sixth columns contain two different measures of inequality: one proposed by Abul Naga and Yalcin (2008) and the other by Allison and Foster (2004). The fifth column calculates the mean happiness levels and the last column provides the total number of observations in each year.

First we discuss the Allison-Foster inequality measure which is based on the mean level of happiness below and above the median which are presented in the third and fourth columns respectively. The difference of the two columns gives us the inequality measure. This can also be shown through the figure below.

[Insert Figure 2]

The plot on the top of the graph shows the mean of the distribution above the median (S_U) and the plot on the lower half shows the mean of the distribution below the median (S_L). If the plots move in the same direction we cannot be sure whether inequality has increased or not. For instance, both S_U and S_L increase from 1972 to 1973, yet we cannot be sure of the changes to inequality since that will depend on the magnitude of increase of S_U and S_L. However, if the plots move in opposite direction, then we can say for sure what will happen to inequality one way or the other. Between 1989 and 1990, S_U has decreased, whereas S_L has increased. In such circumstance, we can say for sure inequality will decrease. If the opposite happens then we can again say for definite that inequality will increase.

For clarity let us consider 1990. From the given information, it is easy to find out that in 1990 there were 7.74 percent reporting Not too happy, 56.53 percent report Pretty Happy and 35.73 percent reporting Very Happy. If we consider a linear scale, say (-1,0,1) then the mean happiness below the median would be given by (0.0774 x -1)/0.5 = -0.1548. Similarly the mean of the upper half will be (0.3573x1)/0.5=0.7146. Thus the difference between them is of 0.8694.

Under the AF measure, with a scale c = (-1,0,1), on an average, the 1970's had happiness inequality of 0.471; it decreased in the 80's to 0.442; in the 90's it further reduced to 0.424. However, since 2000, the average inequality has increased back up and stands at 0.435. The inequality in the 2000's is thus higher than the 1990's but not as high as the 80's or 70's. Thus in some sense, the progress made in reducing happiness inequality through the 1990's had been wiped out in the 2000's.

The Abul Naga and Yalcin (AY) measure which is another measure of inequality for ordinal data based on percentages of the population not in the median category, show a similar trend. From 1972 to 1998 inequality has reduced for each decade, from 0473 in the 1970's to 0.446 in the 80's to 0.427 in the 90's. However in 2000's the inequality has risen to 0.445. Similar to the AF measure, 1974 stands out as the year with the highest inequality and 1985 is the year with the lowest inequality. Interestingly when the scale c = (-1,0,1) is used, on an average the AY (1,1) measures turns out to be half of the AF measure but note that this result may not hold under a different scale.

When the question order corrected data is taken in to consideration, for both the AY(1,1) and the AF measure, inequality increases in all the years where the question order effect is corrected. The overall decadal trends, however, remained the same as before with a steady decline in inequality from the 1970's through to the 1990's accompanied by a rise in inequality in the 2000's. These results are summarised in the table below.

[Insert Table 3b]

Now for both the measures, 1972 is the year with the highest inequality rather than 1974. Under the corrected order, we can see that in 1972, the percentage of the people in the very happy category increases by almost 8 percentage point while both the median category and the lowest category have less number of people in them. 1985, however, remains as the year with the lowest inequality.

Of course, as is clear from the discussion earlier, the levels are completely scale dependent and it is possible that the ranking may be reversed under a different scale. Recall that when the Allison-Foster measure shows one year has more inequality than the other it is just indicative that such might be the case under all scales. To test whether such is the case or not we need to check for S-dominance.

3.2 S- Dominance Criteria

If we compare 1985 with 1974 we find that the share of the population for 1985 and 1974 in the worst category (Not too Happy) are $f^1_{1985}=0.104< f^1_{1974}=0.125$ and the share of the population in the best category (Very Happy) are $f^3_{1985}=0.296<f^3_{1974}=0.383$. In other words, distribution in 1974 has more mass in the extremes (i.e. in the worst and the best category) and hence 1985 S-dominates 1974. What this implies is that for all possible scales, 1985 will have lower inequality than 1974 and we can unambiguously rank 1974 higher in terms of inequality compared to 1985.

As mentioned before, along with S-dominance, if we employ the concept of F-dominance we shall arrive at a complete ranking of the years. For instance if we compare two distributions, say of 1990 and 2006 the share of the population in the worst category was, $f^1_{1990}=0.077$ and $f^1_{2006}=0.105$ respectively. The share of the population in the best category for 1990 and 2006 was $f^3_{1990}=0.357$ and $f^3_{2006}=0.335$. Clearly then $f^1_{1990}< f^1_{2006}$ and $f^3_{1990}> f^3_{2006}$. Although the criteria for S-dominance comparison is not satisfied, still in terms of happiness we can claim that 1990 F-dominates 2006 because a lower share of the population is in the worst category and a greater share in the best category in 1990 as compared to 2006. We can make such comparison for any given pair of years.

For pair wise comparison of all the years we construct a square matrix, with 27 rows and 27 columns with each row and column representing a year from our sample. Each cell of the matrix describes the comparisons between the distributions of the concerned two years using the F-dominance criteria or the S-dominance criteria.

[Insert Table 4a].

To understand the rankings, let us take for example the year 1976. The first cell in the row labelled 1976, compares the happiness distribution in 1976 with 1972 and the F in that cell shows that 1976 F-dominates 1972. In the next cell we compare the distribution of happiness of 1976 with 1973. In this case 1976 S-dominates 1973, that means that compared to 1973, 1976 has a higher percentage of people in the middle category than in the tails, which reflects a decreased spread. Thus 1976 have less inequality in terms of happiness than 1973.

One year that stands out above all else is 1990. It either F-dominates or S-dominates rest of years. The next best year is 1988 with it either F-dominating or S-dominating other years except 1990. This is followed by 1978 and 1989 which are dominated by two and three other years respectively. Both from the welfaristic and the inequality point of view we can unambiguously rank 1990 and 1988 above all others. On the bottom is 1972, which is either F-dominated or S-dominated by all the years and is closely followed by 2008.

Note that in Table 3a, although average inequality in the 90's is lower than average inequality in the 80's, still 1988, 1989 and 1990 dominate all the later years including the 90's and the new millennium. In other words, in terms of happiness, the late eighties were the best years for US and they have not been matched since.

We now explore the dominance relationships for the question order corrected data and the results are presented in Table 4b below.

[Table 4b]

Since the data is corrected for 1972, 1980, 1985, 1986 and 1987, we see changes in the dominance relations only for those years. The most dramatic change is observed for 1972 where for the corrected data, all years except 1974 S-dominate 1972 and 1974 F-dominate 1972, which is exactly the opposite to the uncorrected data. This is because for 1972 under the corrected data there has been almost eight percentage points increase in the Very Happy category and a not so steep (around 4 percentage points) decline in the lowest category of Not So Happy category. Similarly for the other years, we see that they are F-dominated by less number of years. Overall now there are 175 instead of 187 F-dominating relations and 176 instead of 164 S-dominating relations. Although the number of S-dominance relations is now greater than the F-dominance relations compared to the uncorrected data, 1990 still remains the best year in terms of F-dominance and S-dominance, followed by 1988.

4. Decomposition

In this section we are interested in compute group inequalities by gender, race and region. We use the question order uncorrected weighted data for this purpose. In the previous discussions we have seen that both the question order corrected and uncorrected weighted data the broad trends have remained the same. In particular the difference in average inequality for AF and AY inequality measure is around 0.014 and 0.003 respectively between the two data sets. Further, the revised sample weights taking in to account the question order effects are unavailable. Only the percentages of individuals in each of the happiness categories for the whole sample are reported for the five years. Therefore we do not have the question order corrected distribution for different sub-samples based on race, sex or geographical location.

Although S-dominance is a robust criterion for ranking inequality in terms of happiness, here we employ the Allison-Foster measure of inequality for two reasons: first it still is indicative of the S-dominance relationship and second it is fully decomposable. We demonstrate the later part next.

Suppose there are p groups. Each group j has a population n_j. Let the median category, k, be the same across all the groups. Then equation (2) can be written as

$$I_{AF} = 2((\sum_{i>m} c_i.\sum f^i))-(\sum_{i<m} c_i.\sum f^i)+c_m(\sum F^m+F^{m-1} - 1))$$

which implies

$$I_{AF} = \sum((n_j/n)I_j,$$

where I_j stands for the Allison-Foster inequality for group j.

4.1 Gender

It is well established that there are significant differences between the genders when it comes to income and labour market outcomes. In the happiness literature also, there is growing evidence of significant differences between men and women. Clark (1997) finds that very different things make women and men happy. So although women got paid less than men, yet they were more satisfied with their lives then men.

To understand if there are any differences among the genders when it comes to happiness inequality, we decompose the overall happiness inequality to the inequality between male and female. The results are presented in the table below.

[Insert Table 5]

Our sample, over the 27 years, consists of 25,644 women and 21,750 men. For all the years, the second category (Pretty Happy) is the median category for both the genders.

The broad patterns that emerge from the table are that on average women has higher happiness inequality with 0.915 compared to men 0.877. While the lowest inequality for women came in 1985 with 0.795, men had the lowest inequality in 1993 with 0.782. Infact, compared to the other decades, the 90's indicated the lowest happiness inequality on average, for both genders. Both men and women followed the national trend with happiness inequality decreasing from the 70's to the 80's with further falls in the 90's, but inequality came back up again in the current decade. The gap between men and women in terms of inequality, however, has continued to fall through all the four decades and now stands at the lowest.

4.2 Race

In the GSS, race has been classified in three groups: Whites, Blacks, and Others. The sample consisted of, 39,728 of Whites, 6,581 Blacks and 2,009 from Other races. The number of Others surveyed in some years is very small and in some cases the median category is different from those of Blacks and Whites. For both of the later two groups the median category is Very Pretty. We present the results of the inequality decomposition below but our analysis main concentrates on the Blacks and Whites.

[Insert Table 6]

Over the whole period from 1972 to 2008, Blacks on an average had a happiness inequality of 0.880 which is lower than the inequality amongst Whites of 0.901. The lower inequality within Blacks stems from the fact that a greater share of their population is in the median category as compared to the Whites. For instance, in 7 out of 26 years the Blacks have more than 60% of their population in the median category, whereas Whites only had for a year a similar percentage of their population in the median category. The greater dispersion, thus, amongst the Whites is reflected in the higher magnitude of inequality.

There are, however, significant differences in the experience of inequality among the groups. Whilst for Blacks and Whites the inequality was highest in 1974; for the Blacks their lowest levels of inequality came in 1990, whereas for Whites it came in 1985. On an average, Blacks had lower inequality than the Whites, in the 70's, in the 80's and in the 90's. It is only in the current decade that Blacks with 0.453 have more inequality than the Whites with an inequality of 0.438. Interestingly Blacks had the lowest inequality in the 90's and it is also at that time that the difference between the Blacks and the Whites were the largest. Both groups also followed the general trend with happiness decreasing from 70's all the way to the 90's but the inequality had been on the increase recently.

4.3 Regions

There is growing evidence that happiness and the factors that affects it varies across regions (Alesina et al 2004, Graham and Felton 2006). Our primary interest here is whether the distribution of happiness varies across regions with the US. GSS reports happiness levels for 9 zones. We have followed the US Census definitions and collapsed them in to four regions -- the Mid West, the North East, the South and the West. Our sample consists of 9,588 observations from the North East, 12,699 observations from the Mid West, 16,787 observations from the South and 9,244 observations from the West. The following table presents the decomposition of the happiness inequality across different regions.

[Insert Table 7]

On average North East has the lowest happiness inequality over the 27 years and is closely followed by Mid West and the West. The South has the highest inequality with 0.948. This, however, masks considerable variations in inequality across the regions over time. Infact in the 70's Mid West had the highest inequality right after the South. By the 80's, however, it was able to reduce the level of inequality quite drastically. The general pattern of inequality over these fours regions across the last four decades is presented in the table below

[Insert Table 8]

All three regions excluding the Mid West followed the national trend, where inequality decreased through the 80's and the 90's only to increase in the 2000's. For the North East inequality in the 2000's has increased so much that it is now greater that its inequality level in the 80's. The Mid West, on the other hand, continued with the downward trend in inequality even in the 2000's till 2006. However in 2008 there has been dramatic increase in the inequality of happiness in the Mid-West as a result of which it is North-East which has on an average the lowest inequality.

5. Conclusion

Since happiness is an important indicator of subjective well-being, understanding the distribution of happiness across the population is also crucial. The happiness data is of ordinal nature thus the use of standard inequality indices to measure happiness inequality is problematic since the results may be inconsistent under different scales. The justification to use a particular cardinal scale to evaluate the levels of happiness inequality is also adhoc and unsatisfactory.

In this paper, using a methodology developed by Allison and Foster (2004), we calibrate the happiness inequality in the US, from 1972 to 2008. We find that 1990 was the best year both in terms of lower happiness inequality and higher well-being. It was closed followed by 1988 and 1989. In terms of broad trends, happiness inequality decreased, from its highest level in 1970's, through the 1980's and 1990's. Only in the 2000's it has started to rise again. There, are however, considerable variation across gender, race and region.

References

Abul Naga, R., T. Yalcin, 2008. "Inequality measurement for ordered response health data," Journal of Health Economics, 27(6), 1614-1625,

Alesina, A., R. Di Tella, R. MacCulloch, 2004. Inequality and happiness: are Europeans and Americans different? *Journal of Public Economics*, 88, 2009-2042.

Allison, R. A., J. Foster, 2004. Measuring health inequality using qualitative data. *Journal of Health Economics* 23, 505-524.

Clark, A. E., 1997. Job satisfaction and gender: Why are women so happy at work? *Labour Economics*, 4, 341-372.

Di Tella, R., R. MacCulloch, A. Oswald, 2003. The macroeconomics of happiness. *Review of Economics and Statistics*, 85, 809-827.

Graham, C., A. Felton, 2006. Inequality and happiness: Insights from Latin America, *Journal of Economic Inequality*, 4, 107-122.

Kalmijn, W., R. Veenhoven, 2005. Measuring inequality of happiness in nations: In search of Proper Statistics. *Journal of Happiness Studies*, 6, 351-355.

Ng, Y-K, 1996. Happiness surveys: Some comparability issues and an exploratory survey based on just perceivable increments. *Social Indicators Research*, 38, 1-27.

Ng, Y-K, 1997. A case for happiness, cardinalism, and interpersonal comparability. *Economic Journal*, 107, 1848-1858

Oswald, A., 1997. Happiness and economic performance. *Economic Journal*, 107, 1815-1831.

Sen, A. 1987. Standard of Living. Oxford University Press. Oxford.

Stevenson, B., J. Wolfers**.** 2008. “Happiness inequality in the United States. *Journal of Legal Studies*, 37(2): S33–79.

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

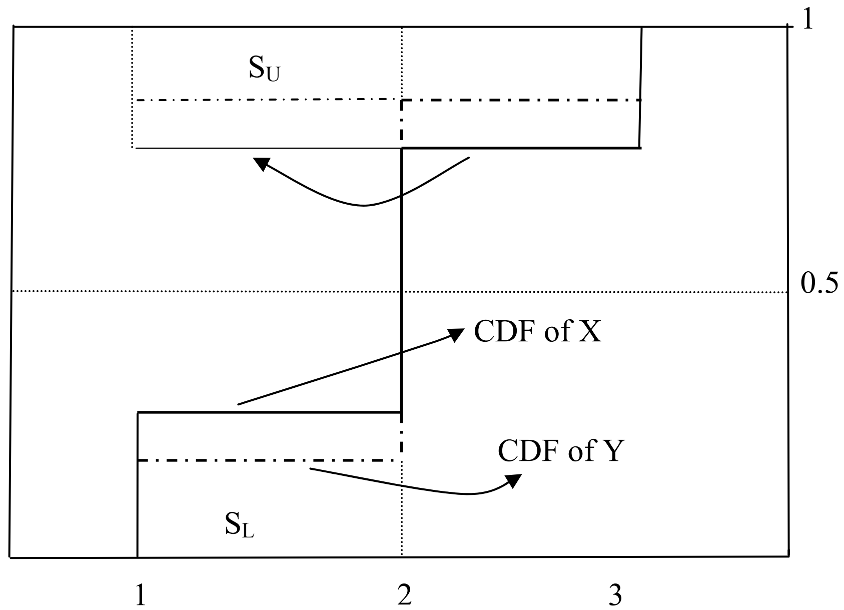

Figure 1: Cumulative distribution and S-dominance for a three category case. Here Y S-dominates X.

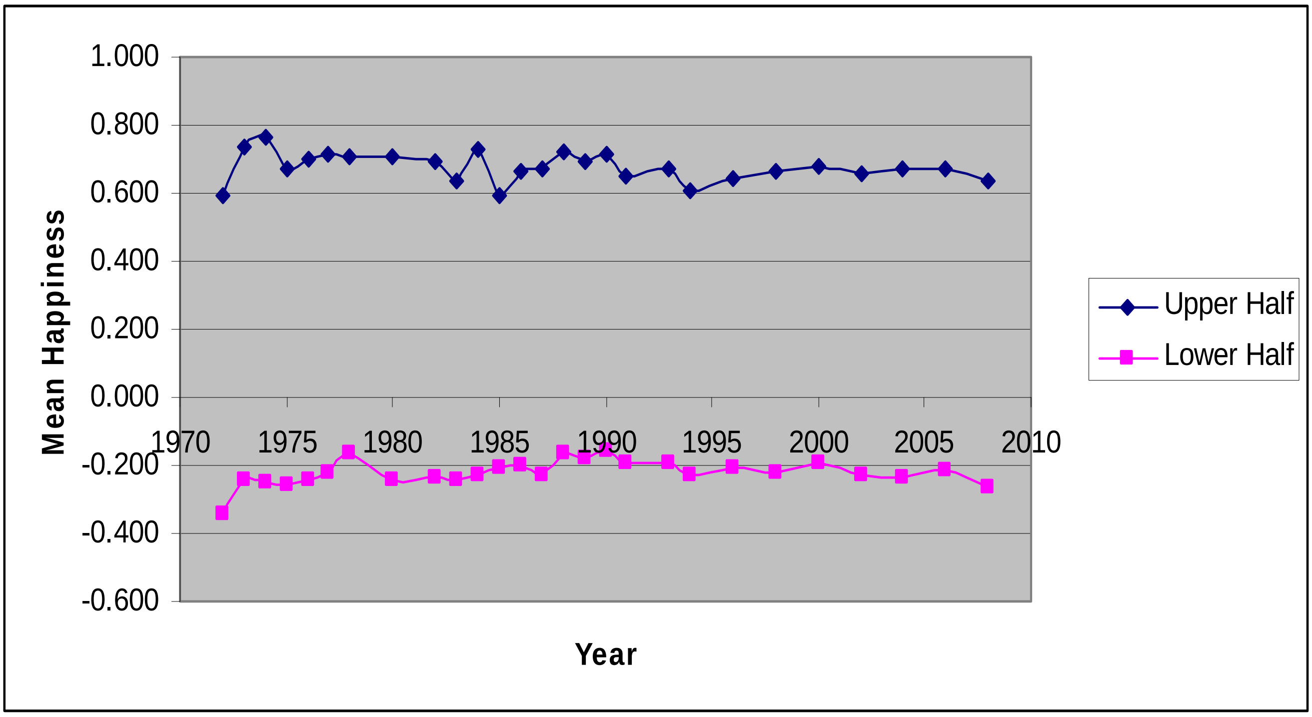

Figure 2: The lower connected lines shows the mean of the lower half of the distribution and the upper connected shows the mean of the upper half of the distribution for different years.
The scale is c = (-1, 0, 1).

Table 1: Summary of the Dominance Relations

Share of Population in Lowest Category	Share of Population in Highest Category	Relation between X and Y
$f^1_X > f^1_Y$	$f^3_X > f^3_Y$	Y S-dominates X
$f^1_X > f^1_Y$	$f^3_X < f^3_Y$	Y F-dominates X
$f^1_X < f^1_Y$	$f^3_X > f^3_Y$	X F-dominates Y
$f^1_X < f^1_Y$	$f^3_X < f^3_Y$	X S-dominates Y

Table 2: Distribution of Happiness across categories

	Frequency	Percent	Cumulative Percent
Not Too Happy	15,921	33.59	33.59
Pretty Happy	26,235	55.35	88.95
Very Happy	5,238	11.05	100
Total	47,394	100	

Table 3a: Population shares and median inequality (for weighted data)

	Share of Population in Not too Happy	Share of Population in Pretty Happy	Share of Population in Very Happy	Mean of the Upper Half (S_U)	Mean of the Lower Half (S_L)	AY (1,1)	AF Inequality	Average Happiness	Total Number of Households
1972	17.230	53.047	29.723	0.594	-0.345	0.470	0.939	0.125	1606
1973	12.275	50.929	36.796	0.736	-0.246	0.491	0.981	0.245	1500
1974	12.507	49.185	38.308	0.766	-0.250	0.508	1.016	0.258	1480
1975	12.967	53.626	33.407	0.668	-0.259	0.464	0.927	0.204	1485
1976	12.240	52.914	34.845	0.697	-0.245	0.471	0.942	0.226	1499
1977	11.025	53.249	35.726	0.715	-0.220	0.468	0.935	0.247	1527
1978	8.362	56.183	35.455	0.709	-0.167	0.438	0.876	0.271	1517
1980	12.099	52.623	35.278	0.706	-0.242	0.474	0.948	0.232	1462
1982	11.686	53.495	34.819	0.696	-0.234	0.465	0.930	0.231	1505
1983	12.085	56.236	31.679	0.634	-0.242	0.438	0.875	0.196	1573
1984	11.603	52.089	36.309	0.726	-0.232	0.479	0.958	0.247	1445
1985	10.457	59.898	29.645	0.593	-0.209	0.401	0.802	0.192	1530
1986	10.174	56.597	33.229	0.665	-0.203	0.434	0.868	0.231	1449
1987	11.451	55.005	33.544	0.671	-0.229	0.450	0.900	0.221	1437
1988	8.240	55.692	36.068	0.721	-0.165	0.443	0.886	0.278	1466
1989	8.793	56.737	34.470	0.689	-0.176	0.433	0.865	0.257	1526
1990	7.740	56.528	35.731	0.715	-0.155	0.435	0.869	0.280	1361
1991	9.485	58.004	32.510	0.650	-0.190	0.420	0.840	0.230	1504
1993	9.736	56.865	33.400	0.668	-0.195	0.431	0.863	0.237	1601
1994	11.287	58.215	30.498	0.610	-0.226	0.418	0.836	0.192	2977
1996	10.517	57.366	32.117	0.642	-0.210	0.426	0.853	0.216	2885
1998	10.896	55.852	33.252	0.665	-0.218	0.441	0.883	0.224	2806
2000	9.644	56.436	33.921	0.678	-0.193	0.436	0.871	0.243	2777
2002	11.284	55.846	32.870	0.657	-0.226	0.442	0.883	0.216	1369
2004	11.694	54.716	33.591	0.672	-0.234	0.453	0.906	0.219	1337
2006	10.553	55.903	33.545	0.671	-0.211	0.441	0.882	0.230	2828
2008	13.287	54.749	31.964	0.639	-0.266	0.453	0.905	0.187	1942
Average	11.086	55.111	33.804	0.676	-0.222	0.449	0.898	0.227	

Note: The scale used is c=(-1,0,1): -1 for Not Too Happy, 0 for Pretty Happy and 1for Very Happy

Table 3b: Population shares and median inequality (for question order corrected data)

	Share of Population in Not too Happy	Share of Population in Pretty Happy	Share of Population in Very Happy	Mean of the Upper Half (S_U)	Mean of the Lower Half (S_L)	AY (1,1)	AF Inequality	Average Happiness	Total Number of Households
1972	13.600	49.100	37.300	0.746	-0.272	0.509	1.018	0.237	1606
1973	12.275	50.929	36.796	0.736	-0.246	0.491	0.981	0.245	1500
1974	12.507	49.185	38.308	0.766	-0.250	0.508	1.016	0.258	1480
1975	12.967	53.626	33.407	0.668	-0.259	0.464	0.927	0.204	1485
1976	12.240	52.914	34.845	0.697	-0.245	0.471	0.942	0.226	1499
1977	11.025	53.249	35.726	0.715	-0.220	0.468	0.935	0.247	1527
1978	8.362	56.183	35.455	0.709	-0.167	0.438	0.876	0.271	1517
1980	11.600	52.000	36.400	0.728	-0.232	0.480	0.960	0.248	1462
1982	11.686	53.495	34.819	0.696	-0.234	0.465	0.930	0.231	1505
1983	12.085	56.236	31.679	0.634	-0.242	0.438	0.875	0.196	1573
1984	11.603	52.089	36.309	0.726	-0.232	0.479	0.958	0.247	1445
1985	8.600	58.400	33.100	0.662	-0.172	0.416	0.834	0.245	1530
1986	9.200	55.800	35.000	0.700	-0.184	0.442	0.884	0.258	1449
1987	9.700	53.300	37.000	0.740	-0.194	0.467	0.934	0.273	1437
1988	8.240	55.692	36.068	0.721	-0.165	0.443	0.886	0.278	1466
1989	8.793	56.737	34.470	0.689	-0.176	0.433	0.865	0.257	1526
1990	7.740	56.528	35.731	0.715	-0.155	0.435	0.869	0.280	1361
1991	9.485	58.004	32.510	0.650	-0.190	0.420	0.840	0.230	1504
1993	9.736	56.865	33.400	0.668	-0.195	0.431	0.863	0.237	1601
1994	11.287	58.215	30.498	0.610	-0.226	0.418	0.836	0.192	2977
1996	10.517	57.366	32.117	0.642	-0.210	0.426	0.853	0.216	2885
1998	10.896	55.852	33.252	0.665	-0.218	0.441	0.883	0.224	2806
2000	9.644	56.436	33.921	0.678	-0.193	0.436	0.871	0.243	2777
2002	11.284	55.846	32.870	0.657	-0.226	0.442	0.883	0.216	1369
2004	11.694	54.716	33.591	0.672	-0.234	0.453	0.906	0.219	1337
2006	10.553	55.903	33.545	0.671	-0.211	0.441	0.882	0.230	2828
2008	13.287	54.749	31.964	0.639	-0.266	0.453	0.905	0.187	1942
Average	10.763	54.793	34.447	0.689	-0.215	0.452	0.904	0.237	

Note: The scale used is c=(-1,0,1): -1 for Not Too Happy, 0 for Pretty Happy and 1for Very Happy

Table 4a: First Order Dominance and S-Dominance for Happiness over 1972-2008 (for question order uncorrected weighted data)

	72	73	74	75	76	77	78	80	82	83	84	85	86	87	88	89	90	91	93	94	96	98	00	02	04	06	08	F	S
72																												0	0
73	F		S	F																							F	3	1
74	F			F																							F	3	0
75	F																										F	2	0
76	F	S	S	F																							F	3	2
77	F	S	S	F	F			F	F	F	S			F						F				F	F		F	11	3
78	F	S	S	F	F	S		F	F	F	S	F	F	F		F		F	F	F	F	F	F	F	F	F	F	20	4
80	F	S	S	F	F																						F	4	2
82	F	S	S	F	S			S		F															F		F	5	4
83	F	S	S	S	S			S																			S	1	6
84	F	S	S	F	F			F	F	F															F		F	8	2
85	S	S	S	S	S	S		S	S	S	S			S						S	S	S		S	S	S	S	0	17
86	F	S	S	S	S	S		S	S	F	S	F		S						F	F	S		F	S	S	F	7	12
87	F	S	S	F	S			S	S	F	S														S		F	4	7
88	F	S	S	F	F	F	F	F	F	F	S	F	F	F		F		F	F	F	F	F	F	F	F	F	F	22	3
89	F	S	S	F	S	S		S	S	F	S	F	F	F				F	F	F	F	F	F	F	F	F	F	16	7
90	F	S	S	F	F	F	F	F	F	F	S	F	F	F	S	F		F	F	F	F	F	F	F	F	F	F	22	4
91	F	S	S	S	S	S		S	S	F	S	F	S	S					S	F	F	S	S	S	S	S	F	6	16
93	F	S	S	S	S	S		S	S	F	S	F	F	S						F	F	F		F	S	S	F	9	11
94	F	S	S	S	S			S	S	S	S			S											S		S	1	11
96	F	S	S	S	S	S		S	S	F	S			S						F		S		S	S	S	F	4	13
98	F	S	S	S	S	S		S	S	F	S			S						F				F	S		F	5	10
00	F	S	S	F	S	S		S	S	F	S	F	F	F					F	F	F	F		F	F	F	F	14	7
02	F	S	S	S	S			S	S	F	S			S						F					S		F	4	9
04	F	S	S	F	S			S		F																	F	4	4
06	F	S	S	F	S	S		S	S	F	S			F						F		F		F	S		F	8	8
08	F																											1	0
F	25	0	0	15	6	2	2	5	5	17	0	8	6	7	0	3	0	4	5	13	8	7	4	10	8	5	22	187	
S	1	22	23	9	15	10	0	15	12	2	16	0	1	8	1	0	0	0	1	1	1	4	1	3	10	5	3		164

Note: S represents the S-dominance relation and F represents the F-dominance relation.

Table 4b: First Order Dominance and S-Dominance for Happiness over 1972-2008 (for question order corrected wreighted data)

	72	73	74	75	76	77	78	80	82	83	84	85	86	87	88	89	90	91	93	94	96	98	00	02	04	06	08	F	S
72																												0	0
73	S		S	F																							F	2	2
74	F			F																							F	3	0
75	S																										F	1	1
76	S	S	S	F																							F	2	3
77	S	S	S	F	F			S	F	F	S									F				F	F		F	8	5
78	S	S	S	F	F	S		S	F	F	S	F	F	S		F		F	F	F	F	F	F	F	F	F	F	17	7
80	S	S	S	F	F				F	F	F														F		F	7	3
82	S	S	S	F	S					F															F		F	4	4
83	S	S	S	S	S																						S	0	6
84	S	S	S	F	F				F	F															F		F	6	3
85	S	S	S	S	S	S		S	S	F	S		S	S		S		F	S	F	F	S	S	F	S	S	F	6	17
86	S	S	S	F	F	S		S	F	F	S			S				F	F	F	F	F	F	F	F	F	F	14	7
87	S	F	S	F	F	F		F	F	F	F								F	F	F	F		F	F	F	F	16	2
88	S	S	S	F	F	F	F	S	F	F	S	F	F	S		F		F	F	F	F	F	F	F	F	F	F	19	6
89	S	S	S	F	S	S		S	S	F	S		S	S				F	F	F	F	F	F	F	F	F	F	12	10
90	S	S	S	F	F	F	F	S	F	F	S	F	F	S	S	F		F	F	F	F	F	F	F	F	F	F	19	7
91	S	S	S	S	S	S		S	S	F	S			S					S	F	F	S	S	S	S	S	F	4	16
93	S	S	S	S	S	S		S	S	F	S									F	F	F		F	S	S	F	6	11
94	S	S	S	S	S			S	S	S	S														S		S	0	11
96	S	S	S	S	S	S		S	S	F	S									F		S		S	S	S	F	3	13
98	S	S	S	S	S	S		S	S	F	S									F				F	S		F	4	10
00	S	S	S	F	S	S		S	S	F	S			S					F	F	F	F		F	F	F	F	10	9
02	S	S	S	S	S			S	S	F	S									F					S		F	3	9
04	S	S	S	F	S					F																	F	3	4
06	S	S	S	F	S	S		S	S	F	S									F		F		F	S		F	6	9
08	S																											0	1
F	1	1	0	16	8	3	2	1	8	19	2	3	3	0	0	3	0	6	7	15	10	9	5	12	11	7	23	175	
S	25	21	23	8	13	10	0	15	10	1	15	0	2	8	1	1	0	0	2	0	0	3	2	2	8	4	2		176

Note: S represents the S-dominance relation and F represents the F-dominance relation.

	Table 5: Decomposition of Median Inequality by Gender					
	Inequality Women	N_W	Inequality Male	N_M	Inequality Overall	N_T
1972	0.959	758	0.921	848	0.939	1606
1973	1.027	782	0.931	718	0.981	1500
1974	1.060	763	0.970	717	1.016	1480
1975	0.967	794	0.882	691	0.927	1485
1976	0.965	803	0.915	696	0.942	1499
1977	1.007	816	0.853	711	0.935	1527
1978	0.867	859	0.889	658	0.876	1517
1980	0.985	809	0.901	653	0.948	1462
1982	0.959	832	0.895	673	0.930	1505
1983	0.874	871	0.877	702	0.875	1573
1984	0.983	834	0.924	611	0.958	1445
1985	0.795	809	0.810	721	0.802	1530
1986	0.878	809	0.855	640	0.868	1449
1987	0.896	788	0.904	649	0.900	1437
1988	0.903	805	0.865	661	0.886	1466
1989	0.879	836	0.848	690	0.865	1526
1990	0.867	742	0.872	619	0.869	1361
1991	0.850	837	0.827	667	0.840	1504
1993	0.927	889	0.782	712	0.863	1601
1994	0.826	1623	0.847	1354	0.836	2977
1996	0.847	1532	0.860	1353	0.853	2885
1998	0.912	1545	0.847	1261	0.883	2806
2000	0.882	1518	0.858	1259	0.871	2777
2002	0.886	696	0.880	673	0.883	1369
2004	0.890	699	0.923	638	0.906	1337
2006	0.896	1558	0.865	1270	0.882	2828
2008	0.927	1037	0.880	905	0.905	1942
Average	0.915		0.877		0.898	

Note: N_W refers to number of women in each year; N_M refers to number of men in each year; N_T refers to the total number of households in the data in each year. Scale c= (-1,0,1)

Table 6: Decomposition of Median Inequality by Race

	Inequality Whites	N_w	Inequality Blacks	N_B	Inequality Others	N_{OTH}	Inequality Overall	N_T
1972	0.937	1330	0.948	271	0.909	5	0.939	1606
1973	0.993	1302	0.933	184	0.581	14	0.981	1500
1974	1.007	1312	1.066	162	0.857	7	1.012	1481
1975	0.936	1317	0.861	165	1.000	4	0.927	1486
1976	0.947	1361	0.938	126	0.333	12	0.942	1499
1977	0.929	1347	1.006	165	0.667	15	0.935	1527
1978	0.886	1346	0.817	155	0.666	17	0.876	1518
1980	0.949	1308	0.961	145	0.526	10	0.947	1463
1982	0.926	1323	0.949	153	1.018	28	0.930	1504
1983	0.876	1406	0.861	150	0.941	17	0.875	1573
1984	0.957	1233	0.906	158	1.146	53	0.958	1444
1985	0.792	1330	0.887	151	0.813	50	0.802	1531
1986	0.892	1233	0.680	173	0.941	43	0.868	1449
1987	0.911	1207	0.759	172	1.086	59	0.900	1438
1988	0.892	1224	0.888	179	0.773	63	0.886	1466
1989	0.881	1309	0.801	146	0.705	71	0.865	1526
1990	0.879	1141	0.708	153	1.072	67	0.870	1361
1991	0.835	1249	0.848	203	0.918	52	0.840	1504
1993	0.881	1347	0.738	169	0.820	85	0.863	1601
1994	0.850	2467	0.781	376	0.726	134	0.836	2977
1996	0.837	2332	0.834	372	1.090	180	0.853	2884
1998	0.886	2228	0.824	378	0.959	201	0.883	2807
2000	0.881	2185	0.854	394	0.801	198	0.871	2777
2002	0.884	1102	1.019	175	0.618	92	0.883	1369
2004	0.906	1061	1.007	147	0.787	128	0.906	1336
2006	0.894	2135	0.879	363	0.810	330	0.882	2828
2008	0.887	1519	1.017	264	0.887	159	0.905	1942
Average	0.901		0.880		0.832		0.898	

Note: N_{OTH} refers to number of households in Other Races; N_B refers to number of Black households; N_w refers to number of White households in South; N_T refers to the total number of households in the data. Scale c= (-1,0,1)

Table 7: Decomposition of Median Inequality by Region

	Inequality North East	N_{NE}	Inequality Mid West	N_{MW}	Inequality South	N_S	Inequality West	N_W	Inequality Overall	N_T
1972	0.991	402	0.853	444	1.013	486	0.872	274	0.939	1606
1973	0.873	357	0.948	424	1.093	476	0.979	244	0.981	1501
1974	0.981	343	0.986	431	1.120	468	0.918	238	1.016	1480
1975	0.900	327	0.901	443	1.006	506	0.835	209	0.927	1485
1976	0.895	345	1.041	424	0.895	486	0.928	244	0.942	1499
1977	0.933	310	0.883	459	0.995	507	0.912	251	0.935	1527
1978	0.727	342	0.837	450	1.009	489	0.893	236	0.876	1517
1980	0.897	308	0.881	400	1.006	494	0.998	260	0.948	1462
1982	1.048	365	0.841	421	0.963	482	0.841	237	0.930	1505
1983	0.799	354	0.838	467	0.946	482	0.912	270	0.875	1573
1984	0.919	293	0.951	414	0.984	475	0.967	263	0.958	1445
1985	0.775	295	0.793	400	0.841	542	0.769	294	0.802	1531
1986	0.867	310	0.826	386	0.885	467	0.899	285	0.868	1448
1987	0.766	287	0.937	393	0.922	491	0.949	266	0.900	1437
1988	0.876	303	0.880	388	0.941	487	0.813	288	0.886	1466
1989	0.880	311	0.849	393	0.913	520	0.789	303	0.865	1527
1990	0.904	278	0.808	371	0.883	459	0.897	253	0.869	1361
1991	0.796	301	0.759	366	0.914	545	0.847	292	0.840	1504
1993	0.764	293	0.831	417	0.905	551	0.918	340	0.863	1601
1994	0.779	587	0.824	706	0.884	1077	0.819	608	0.836	2978
1996	0.835	562	0.865	657	0.887	1019	0.802	647	0.853	2885
1998	0.799	572	0.868	677	0.909	1002	0.940	554	0.883	2805
2000	0.891	577	0.772	652	0.932	986	0.859	562	0.871	2777
2002	0.919	290	0.821	346	0.931	454	0.845	278	0.883	1368
2004	0.874	232	0.844	309	0.960	515	0.901	281	0.906	1337
2006	0.883	485	0.863	632	0.898	1056	0.875	655	0.882	2828
2008	0.843	339	0.953	422	0.965	696	0.821	485	0.905	1942
Average	0.867		0.869		0.948		0.881		0.898	

Note: N_{NE} refers to number of households in North East; N_{MW} refers to number of households in Mid West; N_S refers to number of households in South; N_W refers to number of households in West; N_T refers to the total number of households in the data. Scale c= (-1,0,1)

Table 8: Average inequality across decades for different regions

	70's	80's	90's	00's	Average
North East	0.900	0.870	0.813	0.882	0.866
Mid West	0.921	0.866	0.826	0.851	0.866
South	1.019	0.933	0.897	0.937	0.947
West	0.905	0.882	0.870	0.860	0.879
Average	0.936	0.888	0.852	0.882	