CENTER FOR LABOR ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY
WORKING PAPER NO. 24

Understanding Changes in the Distribution of Household Incomes in New Zealand Between 1983-86 and 1995-98

Dean Hyslop
Department of Economics
UCLA
&
PCD
NZ Treasury

and

Dave Maré
Labour Market Policy Group
NZ Department of Labour

July 2000

Abstract

This paper presents an analysis of changes in the distribution of gross household income and income inequality over the period 1983 – 1998. The analysis applies a semiparametric approach to study the effects of changes in the distribution of household types, and changes in National Superannuation (old age pension), household socio-demographic attributes and employment outcomes, and in the "economic returns" to such attributes and employment outcomes on the distribution of income, and uses kernel density methods to estimate these effects. This approach provides a visual appreciation of the shape of the income distribution, and is important in understanding how each of these factors affected different parts of the distribution over the period. We also estimate the effects of each of these factors on changes in various summary measures of inequality over the period. The results find that changes in household structure (particularly the declining proportion of two-parent families), attributes, and employment outcomes each contribute to the observed increase in inequality, while the changes in returns are estimated to reduce the level of inequality. Collectively these factors account for about 50 percent of the observed increase, depending on the measure of inequality used. The results confirm other research findings that the changes were concentrated during the late 1980s.

We thank David Card, Ken Chay, Ron Crawford, John DiNardo, Brian Easton, Lesley Haines, Benedikte Jensen, Jas McKenzie, John Scott, and seminar participants at the UC Berkeley's labor lunch, Victoria University of Wellington's REF seminar, and the NZ Treasury for helpful comments and discussions, and Matthew Bell and Ivan Tuckwell for expert assistance with numerous data issues related to the HES survey. Hyslop is grateful for the hospitality and support of the Center for Labor Economics at UC Berkeley. Any views expressed are those of the authors and do not purport to represent those of the Treasury or Department of Labour, and all mistakes remain the sole responsibility of the authors.

I: Introduction

A recent literature has documented dramatic increases in the degree of income inequality in New Zealand since the early 1980s.[1] This literature has most typically focused on the Gini coefficient as a summary measure of income inequality, and documented increasing inequality across a broad range of alternative income measures, including both individual and family/household incomes, measures of market, gross and disposable income, and measures of equivalised and non-equivalised income.

Although the rise in inequality has been well documented, the reasons for the increase are less well understood. The period of increasing inequality coincides with a period of dramatic economic and social policy reform in New Zealand, which naturally has lead to considerable interest in possible links between the two, and several "reform" related hypotheses have been proposed to explain the rising inequality. These include the effects of trade liberalisation which have affected domestic industry, hence employment and income; the effects of labour market reform, particularly the Employment Contracts Act (ECA), which adversely affected the bargaining position of workers, hence their employment and income; and social policy reform, which reduced the generosity of income support for some welfare beneficiaries[2]. However, there are a number of alternative mechanisms by which the distribution of income may be widening and inequality increasing, independently of policy reform. These include changes in household structure across the population, changes in the socio-demographic characteristics of households, shifts in the distribution of employment outcomes within households, and changes in the relative income distribution conditional on employment and attributes.

The objective of this paper is to investigate the changes in the distribution of household incomes in New Zealand between 1983 and 1998. The analysis focuses on two issues that distinguish it from

[1] For example, Statistics New Zealand (1999), Podder and Chatterjee (1998), Martin (1998). Easton (1996) provides a recent evaluation of longer term trends in income inequality in New Zealand, with a particular focus on the post-1984 period. Dixon (1996, 1998) has examined changes in earnings and labour market outcomes of individuals over this period. The increase in inequality in New Zealand has been large by comparison with other developed countries' experiences. For example, the Gini coefficient on equivalised disposable income in New Zealand increased from 0.27 in 1982 to 0.33 in 1996; this was similar to the increase in the UK (from 0.28 in 1981 to 0.33 in 1996), and substantially larger than the US (0.34 in 1984 and 1995), and other OECD countries – see Figure 7.2 in Statistics New Zealand (1999).

previous literature. First, and most importantly, we focus on the entire distribution of income, rather than concentrating on the Gini coefficient or some other summary measure of inequality. The distribution of income is the underlying basis for the Gini and other summary measures of inequality that, by themselves, may be relatively uninformative regarding how the income distribution has changed. In addition, although summary measures provide an index of inequality they may be quite sensitive to changes in specific areas of the income distribution. For these reasons, the analysis of changes in the entire distribution of income in different periods enables a better appreciation of where changes in the overall distribution occurred.

Second, we adapt a semiparametric procedure recently developed by DiNardo, Fortin and Lemieux (1996) to focus on how changes in household structure and other sets of factors affect both the overall distribution of household income. This procedure enables the estimation of suitable "counterfactual" distributions that provides a clear visual sense of the impact of various sets of explanatory factors. In addition to this visual appreciation of how the explanatory factors affect specific points in the distribution, the counterfactual distributions can be used to analyse the impact of the explanatory factors on alternative summary measures of inequality over the period. We focus on the effects of changes in four sets of explanatory factors: first, the effects of changes in household structure, defined according to the presence of children, and the numbers and ages of adults in the household; second, the effects of changes in the socio-demographic attributes of households within each type of household; third, the effects of changes in the employment outcomes of households within each household type; and fourth, the effects of changes in the economic "returns" to the various socio-demographic attributes of the households. In addition to these factors, we also control for changes in the statutory rate of National Superannuation (old-age pension) over the period, as this appears to have a salient role in two regions of the overall distribution of household income.

This work provides an initial step in examining how much of the change in the income distribution and inequality may be driven by recent reforms and how much is the result of secular social

[2] The chapters in Silverstone et al (1996) provide analyses of the effects of a range of policy and economic reforms.

and demographic trends. Some factors that influence the distribution of income, such as the household age structure, are clearly independent of policy reform. However, other factors, such as the rise in the incidence of sole-parent families may be the result of either secular trends and/or may be due to economic and social policy reform: for example, the increasing trend in marital dissolution predates the recent reform period.[3] The present research provides an analysis of how much of the observed change in the distribution of income and/or inequality may be associated with these broad sets of observable factors. Also, although the analysis of the explanatory factors is sequential in nature, so that the results depend on the ordering of the explanatory factors, we believe the analysis provides a useful contribution towards understanding the impacts of some of the important correlates of the change in income distribution over this period.

The remainder of the paper is organised as follows. In the next section we discuss the data to be used in the analysis, and present a description of the trends in aggregate inequality measures and possible correlates of these trends. In section III, we introduce the analytical framework that is used to construct the various counterfactual income distributions, and illustrate its use with an application to broad changes in household structure across the population. This analysis concentrates on changes between the three years at the beginning of the sample period (1983-86) and the three years at the end (1995-98). Section IV extends the analysis to consider the effects of changes in National Superannuation rates, socio-demographic attributes and employment outcomes of households, and the economic returns to household attributes, which may play important roles in understanding changes in the distribution of income over the period. In section V we explain how the counterfactual income distributions constructed in sections III and IV can be used to estimate alternative summary measures of inequality, and we use these to decompose the change in inequality over the period into the effects of the various factors we examine. As most of the observed changes in the income distribution were concentrated in the late 1980s, a period of recession in New Zealand, to examine the robustness of the results, in section VI we analyse the changes in the distribution of income over the two subperiods, 1983-86 to 1989-92 and 1989-92 to 1995-98. In

[3] See Davey (1998), Statistics New Zealand (1998)

section VII, as a way to provide a partial-equivalisation across the population of households, we summarise the results of changes in inequality over the full period by household structure. We conclude the paper with a discussion of the results and caveats in section VIII.

The analysis finds that changes in household structure can account for between 10 percent and one-third of the observed changes in the household income distribution and inequality, depending on the specific measure used. In addition, we find that changing sociodemographic attributes of households can account for a similar fraction of the observed changes. Somewhat surprisingly, we find the substantial changes in employment outcomes over the period had relatively modest effects on overall income inequality; however, these changes did have a larger effect on inequality measured at the household-type level. Although changes in National Superannuation play a prominent role in localised changes in the distribution, this factor contributes relatively little to changes in broad measures of inequality, due in part to offsetting changes for singles and couples. Finally, we find no systematic effects of changes in economic returns to attributes on the household income distribution and inequality.

II: Data and Descriptive Analysis

As a backdrop to the analysis that follows in this paper, we begin by describing the data that we use, and presenting an overview of the trends in household income inequality and other factors of interest over the sample period. The data come from Statistics New Zealand's Household Economic Surveys (HES) over the period 1983 – 1998.[4] The HES is a household-based survey, which samples approximately 3,000 households per year, and the HES-year runs from April to March. For the first three years of the period (1983/84-1985/86), the sample frame used for the HES was a simple random sample of households; for the later years (1986/87 onwards), a stratified random sample of households was drawn in each year. Throughout the analysis the data is weighted using the HES sampling weights.

The HES collects information on the household structure, the socio-demographic characteristics and relationships of individuals in the household, together with income from various sources and some

basic labour market information on individuals, and also household expenditure data on various types of goods and services.

Our primary focus of interest in this paper is the measure of total household income from all sources. There is no single "correct" measure of income to use, and each alternative has its advantages and disadvantages, which depend to some extent on the objectives of the analysis. We concentrate on total household income for three principal reasons. First, as the "family" is the basic unit within which the welfare of individuals in general and children in particular is assessed, we prefer a measure of family (or household) income to individual income. Also, because the "family" as distinct from the household is nebulous, we concentrate on the better-defined empirical measure of household income. To the extent that unrelated individuals share the same household (e.g. flatmates), this will tend to overstate the resources available to individuals. On the other hand, to the extent that individuals receive support from outside the household (e.g. students living away from home), this will tend to understate resources available to individuals. Second, as actual income is an empirically meaningful and measurable concept we prefer to use this instead of an "equivalised" measure, which adjusts income according to the size and composition of the household.[5] Although we don't formally equivalise household income across the population, the analysis below is conducted separately for different types of household composition, and thus provides a partial equivalisation analysis. Finally, the dramatic trends in inequality over the period of interest are common to alternative measures of income, which leads us to believe that the choice of total household income should not be crucial for the qualitative nature of the results.[6]

---

[4] The HES was formerly known as the Household Expenditure and Income Survey (HEIS).

[5] Various equivalisations have been proposed. For example, if Y is total household income, then one general approach to equivalise income is $E = Y/(N_A + \kappa N_C)^{\sigma}$, where N_A and N_C are the numbers of adults and children in the household respectively, and σ captures the economies of scale associated with household size. Varying the choice of κ and σ results in alternative equivalisation schemes. One interpretation of unadjusted household income is that $\sigma=0$ and there are infinite economies of scale.

[6] For example, Podder and Chatterjee (1998) report the Gini coefficient for an equivalised measure of gross income increased 14 percent from 0.353 in 1983/84 to 0.404 in 1995/96 using HES data; Statistics New Zealand (1999) reports the Gini coefficient for equivalised market income increased 20 percent from 0.394 in 1986 to 0.471 in 1996, and the Gini for equivalised disposable income increased 27 percent from 0.254 in 1986 to 0.322 in 1996 using Census data. Also, Dixon (1998) reports the Gini coefficient in weekly earnings of full-time employees increased 20 percent from 0.23 in 1984 to 0.28 in 1997 using HES data.

Figure 1 describes the trends in the mean, median and Gini coefficient of total household income from the HES for each year from 1983/84-1997/98, with each series indexed to 100 in 1983/84. This figure suggests some interesting relationships between the level and dispersion in incomes over the period. First, the level, as measured by mean household income, increased about 5 percent between 1983/84 and 1989/90, then fell approximately 10 percent in the first half of the 1990s before rising again for a (net) 5 percent increase over the full period. In contrast, median incomes remained roughly static until 1988, and then fell about 15 percent though until 1994, before rising again and finishing with a net loss of about 5 percent over the period. These differences suggest that lower and middle incomes fell modestly over the period, while high incomes rose sufficiently to more than offset this loss. Second, the Gini increased 20 percent between 1983/84 and 1990/91, and then fluctuated around this level throughout the 1990s. The trend in the Gini roughly tracks the relative difference between mean and median incomes. In fact, the rapid increase in the Gini that occurred between 1988 and 1991, corresponds to a period of considerable divergence between mean and median income.

We next consider a parsimonious set of six household "types" to capture differences in the number of adults, the presence of children, and the life cycle characteristics across households. Specifically, we distinguish between single and multiple adult households, between households with and without children and, for households without children, between households with adults under and over 60.[7] Figure 2a describes the trends in the sample fractions of households in each of these household types. The most salient changes in household composition are the decline in the fraction of households with multiple adults and children (i.e. essentially "standard" two-parent families), and the increasing fractions of single adult households both with and without children. That is, the fraction of multiple adult households with children fell from 37 percent in 1983/84 to 30 percent in 1997/98, while the fractions of

[7] The age criterion is intended to distinguish between predominantly "retired" and non-retired households. For multiple adult households, we have grouped the household into "over 60" if the age of the eldest adult is over 60 and either (i) all adults are over 50; or (ii) the fraction of total household income from National Superannuation is at least 50 percent. Otherwise we have classed the household as "under 60". The eligibility age for National Superannuation was 60 until 1991, and has been increasing at the rate of ½ year per year since then. Although this has affected the retirement behaviour of the those in their low-60s (e.g., see Coleman and Hansen, 1996, Frame, 1999, and also figure

sole parent households increased from 3 to 6 percent, and single adult households increased from 18 to 21 percent over the period. As with the Gini coefficient, most of these changes occurred in the late 1980s. There was very little systematic change in the fractions of the multiple adult (without children) households over the period. To the extent that the distribution of income varies with household structure, these shifts in the distribution of households over the period suggest that they may have an impact on the overall distribution of household income and hence the level of inequality.

A second possible cause of the change in household income inequality that we consider here concerns changes in the employment outcomes of individuals and how this might be correlated within households. Figure 2b presents the trends in the fraction of adults employed fulltime separately for each of the six household types over the 15 year sample period. In each case there is a general decline in full time employment during the mid to late 1980s , and some pick-up in employment during the 1990s, although the full time employment rates typically remain 5-10 percent lower at the end of the 1990s than in the early 1980s. Given the importance of labor earnings to household income, and to the extent that (un)employment is not evenly distributed across households, these findings suggest that employment loss may help explain the rise in income inequality.

A third factor that may be related to changes in household income inequality is changes in the socio-demographic attributes both across and within different types of households. In order to explore this issue, we have aggregated the data into 3-year samples and examine the household characteristics for three sub-periods corresponding to the beginning (1983/84-1985/86), the middle (1989/90-1991/92), and the end (1995/96-1997/98) of the period. Table 1 contains the sample means of the characteristics, together with the median household income and Gini coefficient, in these subperiods for each of the six household groups. We have adjusted the nominal incomes reported in the HES to 1999 dollar values using the CPI that excludes the effects of the Goods and Services Tax (GST).[8]

2b below), we maintain a fixed age criteria over the sample period to separate households.

[8] A 10% rate of GST was introduced on October 1st 1986, and increased to 12.5% on July 1st 1989. We have used a CPI-exGST series estimated by the Reserve bank of New Zealand. Given the substantial changes in the tax and benefit regimes in the late 1980s, it is not obvious what is the best method to adjust the nominal reported incomes in the HES

Table 1 shows the differences in the level of income across, and trends in relative income inequality within, each of these household types which highlights the importance of understanding how changes within household types affect aggregate changes in household incomes. As expected multiple adult households have higher incomes than single adult households, non-retired ("under 60") households have higher incomes than retired households, and households without children have higher incomes than those with children. Also, although the level of income inequality within each household type, as measured by the Gini coefficient, was similar early in the period (ranging between 0.27 and 0.30 for 5 of the groups, and 0.33 for single adult households), the change in inequality over the period varies substantially by household type. For example, although the increase in inequality in the "under 60" households without children and the multiple adult with children households was similar to that for all households, income inequality actually decreased within the single adult with children household group, and showed more modest increases in the "over 60" households. Other trends worth noting in table 1 include the ageing of adults over the period which appears in all household types except for the single adult with children households, and the fractions of adults with University level qualifications which have increased quite strongly over the period. In addition, while the level of full-time employment fell quite dramatically during the 1980s and only partially recovered in the 1990s, the incidence of part-time employment has also grown substantially. Each of these factors may influence the distribution of income over the period.

III: Analysis of Changes in the Household Income Distribution between 1983-86 and 1995-98

The descriptive analysis of changes in the Gini coefficient of household income inequality, together with the changes in the distribution of household structures, and changes in employment outcomes and attributes of households within 40each household type described above suggest that these factors may provide some contribution to understanding the factors driving the dramatic increases in

---

to constant-price values. We believe that an adjustment excluding GST effects is more suitable for market incomes, while it may be less so for benefit income.

income inequality over this period. In this section we develop an analytical framework to consider the influence of these factors more formally. This analysis uses kernel density estimation techniques to estimate the entire distribution of income. As this requires relatively large samples in order to obtain reliable estimates, we use the aggregated 3-year HES samples described in table 1. Thus, the estimates of the cross-sectional income distribution will be susceptible to changes in the distribution within each three-year period. However, the trends in inequality, household structure and employment outcomes apparent in figures 1 and 2a&b suggest that the bulk of the changes which occurred between 1983 and 1998 occurred between rather than during the three sub-periods we analyse, so that each three-year subperiod is reasonably homogeneous in terms of the factors of interest.

The analytical framework that we adopt uses a semiparametric conditional density estimation technique developed recently by DiNardo, Fortin and Lemieux (1996). This framework has several features. First, it allows an assessment of the entire distribution of household income at a given point in time, and changes in the distribution over time, rather than simply a summary measure of income inequality for the distribution. Second, it allows a sequential decomposition of the overall change in distribution of income into that due to changes in various sets of factors. Third, it enables an assessment of how changes in the various sets of factors examined above affect changes in alternative summary measures of income inequality. In this section, we will discuss the first two of these features and develop the framework used to examine the effects of changes in the distribution of households on the distribution of household income. In the next section we extend this analysis to consider the effects of various sets of factors which may affect the distributions of income within household types. We then use this framework to decompose changes in measures of inequality into the effects of the various sets of factors.

The Distribution of Household Income

We begin by describing the overall distribution of household income from the HES samples.[9] In order to better identify relative changes in incomes we use the (natural) logarithm of income. For

example, if all household incomes increased by 10 percent, then the income distribution would have the same shape and simply be shifted (approximately) 0.1 log-points to the right. In addition, in order to control the length of the tails of the distribution, we have "censored" the income data below at 7.8 log-points (approximately $2,400) and above at 12.5 log-points (approximately $268,000).[10]

Figure 3a presents kernel density estimates of the distributions of household incomes over the two sub-periods 1983-86 and 1995-98.[11] The horizontal axis is on a logarithmic scale, so that the distance from any income level y_0, to $2y_0$ will be the same irrespective of y_o. The solid line represents the smoothed estimate of the distribution of income in 1983-86, while the dashed line represents the estimated distribution in 1995-98. There are several interesting points to note about these distributions. First, both are multi-modal, with three distinct peaks around $12,000 – $15,000, $20,000 – $23,000, and $50,000 – $60,000 respectively. The latter is the main area of concentration of households, while there is a localised concentration of households in the former two ranges.

Second, figure 3a shows there have been dramatic shifts in the income distribution between 1983-86 and 1995-98. This is perhaps even more apparent from figure 3b which shows the estimated change in the density of the income distribution (f(y)) at each real income level y:

$$\Delta\hat{f}(y) = \hat{f}_1(y) - \hat{f}_0(y), \tag{1}$$

where $\hat{f}_0(y)$ and $\hat{f}_1(y)$ denote the estimated density of income in 1983-86 (period "0") and 1995-98 (period "1") respectively. In particular, there has been a large drop in the fraction of households with mid-range incomes (between about $30,000 and $85,000) from 56.9 percent to 48.1 percent. This drop in middle-income households has been matched by a rise in low-income households (with incomes between $15,000 and $30,000) from 21.7 percent to 26.2 percent, and a rise in high-income households (with

---

[9] The appendix contains a detailed description of kernel density estimation, and other related issues.

[10] That is, any household whose log(income) < 7.8 has been changed to 7.8, or log(income) > 12.5 has been changed to 12.5. The fraction of households with left censored income (i.e. log(income)<7.8) remained steady over the sample period at about 1 percent, while the fraction with right censored income (i.e. log(income)>12.5) rose from 0.2 percent in 1983-86 to 0.6 percent in 1995-98.

[11] The appendix provides a brief discussion of kernel density estimation techniques and details of the application to this

income greater than $85,000) from 11.5 percent to 14.8 percent.

A third feature of the income distributions in figure 3a is that there has been an apparent shift in the "spikes" at the low end of the income distribution. In particular, the spike in the distribution around $12,000 (in 1983-86) appears to have shifted to the right by 5-10 percent which, given that median household income fell about 9 percent (see table 4), represents a relative increase in income for households in this area. Also, the spike around $23,000 (in 1983-86) appears to have shifted to the left by 5-10 percent (as well as increasing in magnitude). Finally, note that the small peaks that appear in the tails of the distributions reflect the extent of bottom and top censoring in the data.

A comparison of the income distributions in 1983-86 and 1995-98 shown in figures 3a and 3b gives a broad sense of how income varied across households in each period, and also where changes in the income distribution occurred between the two periods. The principal objective of the remainder of the analysis is to try to identify the impact of observed changes in various factors on the distribution, and how these translate into summary measures of income inequality.

Decomposing Changes in the Distribution of Income

We now turn to the second feature of the analytical framework, which is to provide a decomposition of the overall change in the distribution of household incomes into that attributed to various sets of factors. In this paper, we concentrate on the effects of changes in the distribution of households, changes in the socio-economic attributes of households, changes in the employment outcomes of households, and changes in the economic "returns" to attributes. In this section we detail the decomposition of income changes attributable to changes in the distribution of households, and defer the discussion of the other factors of interest to the next section.

We consider the six discrete "types" of households described in the previous section and defined according to the presence or absence of children, the number of adults and, for households with no children, the age structure of the adults in the household. To examine the effects of changing distribution

---

case.

of households on the distribution of household income, it is helpful to first understand how the overall distribution of income across the population of households is constructed from the distribution of incomes for each household group. To do this, note that the overall distribution of household income is simply the weighted average of each of these sub-distributions, where each sub-distribution is weighted by the fraction of the population of households in that group. That is, if w_{tj} is the weighted fraction of households in type j in period t,[12] and $f_{tj}(y)$ is the probability density of log income y for household type j in period t, then the overall household income distribution can be expressed as

$$f_t(y) = \sum_{j=1}^{6} w_{tj} f_{tj}(y). \qquad (2)$$

In order to illustrate the contributions of the income distributions of the six household types to the overall distribution, we estimate the densities of income for each of the six household groups in 1983-86 and 1995-98 ($f_{tj}(y)$, j=1, …, 6), and weight each by its sample fraction, w_{tj}. Figures 4a and 4b plots these weighted sub-distributions for the 1983–86 and 1995–98 periods respectively, together with the overall distribution of household income (as shown in figure 3a).

There are several important observations to note from figures 4a and 4b. First, as was noted in table 1, the mean (or median) income levels of the different household types are substantially different. Multiple-adult households have predominantly higher incomes than single-adult households; households with children tend to have lower incomes than households without children (i.e. compare the multiple-adult households with and without children, and the single-adult households with and without children); and the over-60s households tend to have lower incomes than the under-60s households. These differences reflect a combination of household size, effects of child rearing, and life-cycle factors.

Second, the relative positions of the income distributions for the individual household types are suggestive of the explanations for the three modes in the overall distribution of income. For example,

---

[12] Note that the weighted fraction of households in household type j is simply the sum of the sampling weights for households in type j: $w_{tj} = \sum_{i=1}^{N_{tj}} \theta_{ti}$, where N_{tj} is the number of type-j households in period t.

although single-over-60s households account for only about 10 percent of all households, they are concentrated in the region of the left-most spike in the overall distribution and account for almost this entire spike. This suggests that this income spike may correspond to National Superannuation (old-age pension) for single people. Similarly, the second income spike (around $20,000 – $23,000) is predominantly due to multiple-over-60s households, which suggests that it corresponds to National Superannuation for married couples.

The households at the main peak in the income distribution (around $50,000 – $60,000) are mainly comprised of multiple-adult households (with and without children). The fall in the fraction of multiple adult households suggests that this is a possible cause of the flattening out of the distribution in this range. The third observation from figures 4a and 4b, which is particularly important for our analysis, is that the relative contributions of the various household-type income sub-distributions changed over the period, reflecting changes in household structure shown in figure 2a. Perhaps the most salient change has been the decline in the fraction of "multiple adult with children" households, together with an increase in the fraction of "single adult with children" households. As a result of this shift from two-parent to sole-parent families, we would expect to see a change in the overall distribution of household incomes. More generally, given the differences in the underlying income distributions for the various household-types, we would expect changes in the distribution of households to cause changes in the overall distribution of household income.

The question is how much of a change would we expect to occur because of changes in the distribution of household-types in the population? The answer to this question depends on how changes in the distribution of households affect different points in the income distributions of the various household-types. For example, there may be quite different implications for the change in the overall distribution of income if the drop in two-parent families is due to a fall in low-income families rather than a fall in high-income families. For this reason, we consider a more restrictive counterfactual distribution which holds constant the income distributions for each household type as they were in 1983-86, and allows only the fraction of household-types to change between 1983-86 and 1995-98. Specifically, we

estimate this distribution by

$$\hat{f}_0^H(y) = \sum_{j=1}^{6} w_{1j}\,\hat{f}_{0j}(y). \tag{3a}$$

This counterfactual distribution involves simply reweighting the 1983-86 income distributions of each household type by the 1995-98 household-type fractions.[13] It is also worth noting that, if the only change that occurred during the period was in the distribution of households in the population, then the income distribution for each household type would be unchanged and the counterfactual distribution for changes in the distribution of households would exactly match the actual distribution of income in 1995-98.

Figure 5a plots this counterfactual distribution together with the actual 1983-86 distribution of income. As the counterfactual distribution simply reweights the underlying distributions for each household-type, the shape of these two distributions is very similar. The counterfactual distribution has somewhat less mass in the middle-income range than the actual distribution and more mass in the lower-income range, reflecting the effects of the shift away from multi-adult households with children towards sole-parent families and households without children over the period. That is, the observed changes in the household structure over the 1980s and 1990s would be expected to cause a downward shift in mass in the distribution of income.

In order to describe how much of the change in the distribution of household income is explained by this change in the distribution of households we compute the difference between the counterfactual and actual 1983-86 distributions:

$$\Delta\hat{f}^H(y) = \hat{f}_0^H(y) - \hat{f}_0(y) = \sum_{j=1}^{6} (w_{1j} - w_{0j})\hat{f}_{0j}(y). \tag{3b}$$

Figure 5b graphs this explained change, together with the estimated total difference between the actual 1983-86 and 1995-98 distributions, described by equation (1). This again shows the effect of the change

[13] Note that this counterfactual construction assumes that the shifts in the distribution of household-types occur randomly across the income distribution for each type. Although this is arguable, it is conditional on no other changes and does involve a reasonably neutral stance on where changes in the distribution occur. In addition, in the next section we allow for changes in the sociodemographic attributes and employment outcomes of households, which will relax the

in household structure would be to shift the distribution of income downwards. Figures 5a and 5b suggest that the changing distribution of household-types provides a partial and important explanation for the observed change in the distribution of household income over the period. For example, the density changes attributable to changing household types account for 21 percent of the total density change.[14] Given that the predicted shift is from the middle-income range to the left hand tail, it is likely to translate into an increase in income inequality. We return to this issue in section V.

IV: Analysis of Changes in Other Factors

In this section, we extend the techniques of constructing counterfactual distributions described above to consider the effects of factors that might affect the distribution of income within household-types. First, given the salient contribution of retired households to the two lower spikes in the income distributions in figures 3a and 3b together with the apparent shifts in these spikes, we examine the effects of changes in the statutory rates of National Superannuation (NS) for retired singles and couples. We then examine the effects of changes in the socio-demographic characteristics of households, changes in the employment outcomes of households, and changes in the economic returns to the socio-demographic factors.[15] We adopt a sequential approach to the analysis examining each of these factors in the order mentioned.

At each stage of the analysis, we construct a suitable counterfactual distribution for the set of factors which conditions on the effects of changes in the factors previously analysed. In the case of the National Superannuation and economic returns counterfactuals this involves adjusting the income level of the household to take account of the changes over the period, while in the case of the socio-demographic

---

strength of the conditioning set.

[14] Specifically, the coefficient from the regression of the density change due to changes in household types on the total density change over the period is 0.21. This measure of the "correlation" between the density change attributable to each factor and the total change in the density has the advantage that it sums to 1 across the set of explanatory (and unexplained) factors that we examine. For this reason, we use this as a simple descriptive measure for accounting purposes in the text, but adopt a more formal measure to compare distributions later in the paper.

[15] For the two "over 60s" household types, because these households are essentially "retired", in the analysis below we only adjust the incomes for National Superannuation changes.

attributes and employment outcomes this involves adjusting the sampling weight of the households. Table 2 summarises the adjustments made to the household income and/or weight for each counterfactual. In the text we discuss the intuition for the analysis and results, and leave the details of the counterfactual constructions to the appendix.

National Superannuation

Given the apparent importance of the income distributions of the "single adult over 60" and "multiple adults over 60" households to the two spikes at the lower end of the overall distribution of household income, and the shift in these spikes over the 1980s and 1990s, we first investigate whether these shifts are essentially due to changes in the National Superannuation rates for singles and married couples.[16] Although, in principle, this analysis could be extended to all welfare benefits we restrict our attention to National Superannuation for two reasons. The contributions of National Superannuation are a salient feature of figures 3a and 3b, and are also straightforward to identify in the data. In contrast, other non-age related welfare benefits play a less salient role, experienced a myriad of different rate changes, and are less easily identified in the data. In addition, the apparent magnitudes of these shifts are approximately equal to the changes in statutory National Superannuation rates for singles and couples over the period (approximately 5.8 percent and -5.4 percent on average respectively).

We construct a counterfactual distribution to take account of changes in the National Superannuation rates, as follows. First, we construct separate indexes for the statutory single-rate and couple-rate between 1983 and 1998 (adjusted for CPI ex-GST changes). Second, for each household in 1983-86, we adjust the National Superannuation component of income to the 1995-98 value, using the single-rate index if the household has a single person aged over 60, and the couple-rate index if the

[16] An alternative hypothesis is that shifts in these spikes are due to an incorrect CPI-adjustment. However, one factor that leads us to believe that the shifting spikes are not simply due to the CPI adjustment that we have made is that these spikes move in opposite directions. That is, the lower-most spike (corresponding to the singles-rate) moves to the right, while the second spike (corresponding to the married couples-rate) moves to the left.

household has two or more members aged over 60.[17] Third, we compute the 1995-98 counterfactual total household income by adding this adjusted National Superannuation to the other actual income received by the household in 1983-86. This provides an estimate of the level of income the household would have received in 1995-98 given their 1983-86 NS and other incomes and the NS statutory-rate changes over the period. Fourth, we obtain kernel density estimates of this counterfactual total household income for each of the household types, using the 1983-86 sampling weights. Finally, the counterfactual income distribution over all households that allows for NS rate changes and changes in the distribution of household types, is obtained by taking the weighted average of these counterfactual distributions using the 1995-98 HH-type weights, as described in the previous section. That is, this counterfactual combines the effects of the changes in household structure and the changes in the NS rates, by replacing the 1983-86 actual distributions with the counterfactual distributions that take account of changes in the NS rate. We denote the estimated counterfactual income distribution over all households

$$\hat{f}_0^{NH}(y) = \sum_{j=1}^{6} w_{1j}\, \hat{f}_{0j}^{N}(y), \tag{4a}$$

and the estimated marginal change in the distribution explained by the statutory changes in NS rates, conditional on changes in household types, is

$$\Delta\hat{f}^{N}(y) = \hat{f}_0^{NH}(y) - \hat{f}_0^{H}(y) = \sum_{j=1}^{6} w_{1j}(\hat{f}_{0j}^{N}(y) - \hat{f}_{0j}(y)) \tag{4b}$$

Figure 6a shows the counterfactual distribution (equation 4a) together with the counterfactual which allows for household type changes shown in figure 5a (equation 3a), while figure 6b shows the change in the income distribution due to NS (equation 4b), together with the total change over the period (equation 1b).

As expected the effects of controlling for changes in National Superannuation are concentrated in

[17] As expected, the vast majority of National Superannuation is paid to "Single Adult over 60" households and "Multiple adult over 60" households with just two members. However, a small fraction of households in the latter group have more than two members over 60 (0.9 percent); similarly a relatively small fraction of households in the other groups also receive NS (4.7 percent). This latter group is largely in the "multiple adults under 60" group.

the lower half of the income distribution, and generate quite noticeable shifts in this region. In particular, these changes cause the left-most spike in the distribution to shift rightwards and the second spike to shift leftwards by about 5 percent in each case. These changes support the hypothesis that the observed shifts in these two spikes are largely due to statutory changes in National Superannuation rates. For example, figure 5b shows that these changes explain most of the distributional changes around these spikes, especially for the lower spike. The density changes attributable to changing National Superannuation rates account for 23 percent of the total density change.

Socio-demographic Attributes

We next consider the effects of changes in the socio-demographic attributes of households on the distribution of income. Changes in household attributes allow us to control for non-random changes in the observable household characteristics across the distribution of household types over time. Changes in the characteristics of households can be expected to lead to changes in the distribution of income across households, irrespective of any changes in the distribution of households. There are at least two ways in which socio-demographics may affect the income levels. First, indicators of human capital such as education levels and experience directly affect individual incomes and hence the level of family or household income. Second, fertility decisions and other life-cycle factors may affect individuals' and households' preferences and labour supply choices and, if so, can be expected to affect the level of household income. For example, if over time families have fewer children and individuals become better educated then we might expect the labour force participation rates of married women to rise and to individuals earning higher incomes and, in turn, each of these to lead to an increase in the household income of these families.

In order to examine the effects of changes in household attributes, we construct a counterfactual distribution which conditions on the changes in household types and NS rates, described above. For each of the four "under 60" household types, we first estimate the relationship of how the household attributes vary between 1983-86 and 1995-98, adopting flexible specifications which vary across each of the six

household types. The attributes we use include the number, age, sex, and ethnic and education structure of adults in the household, together with the numbers of children in various age groups.[18] Second, we use this estimated relationship to adjust the 1983-86 period observations' sampling weights to reflect the change in attributes in 1995-98. This reweighting scheme acts to give greater weight to the 1983-86 households with attributes which are more similar to those in 1995-98 and less weight to households whose attributes are less similar. Third, we construct counterfactual income distributions for each household type using these "reweights" instead of the original HES sample weights, and then construct the counterfactual distribution of income for all households by computing the weighted average of the household type counterfactual distributions using the 1995-98 household type weighted fractions.

In particular, we denote the estimated counterfactual distribution of household income for changing household types, National Superannuation rates and attributes by

$$\hat{f}_0^{XNH}(y) = \sum_{j=1}^{6} w_{1j} \hat{f}_{0j}^{XN}(y), \tag{5a}$$

and the estimated marginal change in the distribution that is explained by the change in attributes is

$$\Delta\hat{f}^{X}(y) = \hat{f}_0^{XNH}(y) - \hat{f}_0^{NH}(y) = \sum_{j=1}^{6} w_{1j}(\hat{f}_{0j}^{XN}(y) - \hat{f}_{0j}^{N}(y)) \tag{5b}$$

Figure 7a shows this counterfactual distribution together with the counterfactual distribution described in figure 6a which allows for changes in the distribution of household types and NS rates. Figure 7b displays the marginal effect of changing attributes on the income distribution together with the total change in the actual income distribution over the period.

Broadly speaking the observed changes in attributes is predicted to cause an upward shift in mass in the distribution of income. For example, there is a (net) drop in the numbers of households in the $20,000 – $65,000 range, and a net gain in the numbers of households with income above $65,000. Figure 6b suggests that the marginal effect of changes in socio-demographic attributes explains perhaps

[18] The specific sets of attributes differ by household type, and are described in the appendix. Appendix tables A1(a) – A1(d) contain the estimation results for these specifications, together with those for the employment outcomes and

20 percent of the total change in income distribution between 1983-86 and 1995-98. As the shift in mass occurs from the middle to the (right hand) tail of the distribution, the changing socio-demographic characteristics of households is likely to cause an increase in income inequality. The density changes attributable to changing household attributes account for 7 percent of the total density change.

Employment Outcomes

The next factor we examine is the employment outcomes of households. Given that the majority of individual and household income comes from labour earnings, obviously the labour supply decisions and employment outcomes of individuals and their households will directly affect the distribution of household income. Furthermore, given the dramatic changes in employment that occurred during the late 1980s, such employment outcomes have potentially important implications for the observed changes in the distribution of income. In order to investigate these issues, we model these employment effects as follows.

First, we construct a fulltime-equivalent (FTE) worker index for each household equal to the number of fulltime employed adults plus one-half of the number of part-time employed adults in the household. Second, for each of the four household-types, we categorise the household's employment outcome into either two (for single adult households) or three (for multiple adult households) discrete groups on the basis of the number of FTE workers in the household. Table 3 describes the employment-outcome grouping for each household-type, together with the distributions of outcomes in the three periods 1983-86, 1989-92, and 1995-98. Broadly speaking, the information in table 3 confirms that there was a drop in employment between 1983-86 and 1989-92, followed by a partial recovery by 1995-98. For example, the fraction of households with at least 2 FTE workers in the "multiple adult with children" households dropped from 41 percent in 1983-86 to 31 percent in 1989-92 before recovering to 32 percent in 1995-98.

Third, for each of the household types, we estimate relationships between the employment

---

economic returns discussed below.

outcomes and household attributes separately for each period, and use the estimated 1983-86 and 1995-98 relationships to predict the employment outcome of each 1983-86 household given its set of attributes. We then further adjust the 1983-86 sampling weights to take account of the changing employment outcomes that occurred between 1983-86 and 1995-98. As in the case of the attributes, this reweighting procedure acts to give more weight to those 1983-86 households which have employment outcomes more similar to the 1995-98 outcomes predicted by the household's attributes, and less weight to those households which have employment outcomes less similar to the predicted 1995-98 outcomes. The counterfactual income distribution for each household type is then estimated using the 1983-86 income data (adjusted for NS changes) and sample weights adjusted for the combined reweighting effects of changing attributes and employment outcomes. The counterfactual income distribution for all households is obtained as the weighted sum of these household type distributions, each weighted by the 1995-98 household-type sample fraction. We denote this estimated counterfactual distribution

$$\hat{f}_0^{EXNH}(y) = \sum_{j=1}^{6} w_{1j}\,\hat{f}_{0j}^{EXN}(y), \tag{6a}$$

and the estimated marginal change in the distribution that is explained by the change in employment outcomes is

$$\Delta\hat{f}^{E}(y) = \hat{f}_0^{EXNH}(y) - \hat{f}_0^{XNH}(y) = \sum_{j=1}^{6} w_{1j}(\hat{f}_{0j}^{EXN}(y) - \hat{f}_{0j}^{XN}(y)) \tag{6b}$$

Figure 8a shows this counterfactual distribution together with the previous counterfactual, which allows for changes in the distribution of household types, NS rates and household attributes, while figure 8b displays the marginal effect of changing employment outcomes on the income distribution together with the total change in the actual income distribution over the period. These figures highlight that the changes in household employment outcomes cause a broad downward shift in mass from the top half to the bottom half of the distribution, reflecting the predominant drop in employment over the sample period. The density changes attributable to changing household employment outcomes account for 20 percent of the total density change.

Economic Returns to Attributes

The final factor that we consider is the effect of changes in the "economic returns" to attributes over the period on the distribution of household incomes. For example, if the returns to education and other measures of skill increased over the period, then this may increase the dispersion in individual income and, if there is positive assortive matching of couples, may exacerbate the dispersion in household incomes. Maani (1999) and Dixon (1998) document an increase in the income and earnings premium associated with qualifications between the mid 1980s and the mid 1990s.

To adjust the distribution of income for changes in the returns to attributes, for each household type, we first estimate specifications for log income (adjusted for NS changes) in terms of the sets of socio-demographic attributes described above. The specifications are estimated separately for the1983-83 and 1995-98 periods.[19] We next use these estimated specifications to predict the change in "returns" that each 1983-86 household would expect to receive between 1983-83 and 1995-98 given its set of attributes, and use this predicted change to adjust the household's 1983-86 log income. We then construct the counterfactual income distribution which replaces the 1983-86 log incomes (adjusted for NS-rate changes) with those that would have resulted if the 1995-98 returns had existed, using the adjusted sample weights which take account of the changes in the distributions of households, attributes and employment outcomes. We denote this estimated counterfactual by:

$$\hat{f}_0^{REXNH}(y) = \sum_{j=1}^{6} w_{1j}\, \hat{f}_{0j}^{REXN}(y), \tag{7a}$$

and the estimated marginal change in the distribution that is explained by the change in economic returns to attributes is

[19] As our objective here is to translate the distribution, and in order to minimise the effect of possible outliers, we use Median (Quantile) regression techniques rather than least squares regression. We also omit from the regression estimation any households that have either left or right censored log income, however we adjust the actual (uncensored) incomes of these households before recensoring the adjusted log income at 7.8 and 12.5. The 1983-86 log income measure we use is that which has already been adjusted for statutory changes in National superannuation, as described above.

$$\Delta\hat{f}^{R}(y) = \hat{f}_0^{REXNH}(y) - \hat{f}_0^{EXNH}(y) = \sum_{j=1}^{6} w_{1j}(\hat{f}_{0j}^{REXN}(y) - \hat{f}_{0j}^{EXN}(y)). \quad (7b)$$

Figure 9a shows this counterfactual distribution together with the previous counterfactual, which allows for changes in the distribution of household types, NS rates, and household attributes and employment outcomes, while figure 9b displays the marginal effect of changes in the returns on the income distribution together with the total change in the actual income distribution over the period.

Figures 9a and 9b suggest there has been little systematic effect of changes in the returns to the socio-demographic characteristics of households over this period on the household income distribution. In fact, the density changes attributable to changing returns account for -5 percent of the total density change, which suggests that the changes predicted by changes in the returns work the actual changes which occurred over the period.

Summary of Explained Changes

Before we move on to consider the implications of this analysis for various summary measures of income inequality, we first summarise the graphical analysis presented in this section. To do this, in figure 10a we graph the counterfactual density given by equation (7a) together with the actual 1983-86 distribution of household income, while in figure 10b we graph the difference between these two densities, together with the total change in density between 1983-86 and 1995-98. Figure 10a shows that these factors tend to generate a noticeable downward shift in mass from the centre of the distribution, as well as most of the shift in the lower spikes, and a small increase in mass in the right hand tail of the distribution. In terms of the observed changes, figure 10b shows that these factors provide a reasonable account of the changes in the lower part of the distribution, but explain less of the changes in the upper region of the distribution. On net, the density changes attributable to combined set of explanatory factors we have examined account for about two-thirds of the total density change.[20]

An alternative way to consider how well these sets of factors explain the distributional changes is

to compare the aggregate counterfactual distribution (equation 7a) with the actual 1995-98 distribution of income, and examine the unexplained changes in density. These are presented in figures 11a and 11b respectively. Again these graphs tend to highlight that, although the factors provide a reasonable explanation of the observed changes in the lower half of the distribution, they are less successful in explaining the changes which occurred in the top half of the distribution. In particular, there has been a significant increase in the fraction of households with income over $100,000 and this change is largely unexplained by the sets of factors that we have examined in this analysis.

One possible explanation for the increasing households in the higher income ranges that we have not been able to examine here may be related to changes in the intensity of employment over the period. For example, Dixon (1998) finds that there has been a large increase in the numbers of hours worked by employed workers in the last 10-15 years, and that the increase is positively correlated with wages. This suggests that this factor might account for at least part of the increase in the fraction of high-income households. This would be true to the extent these changes are not correlated with the observable socio-demographic characteristics of households; otherwise they would show up as increasing returns to attributes. The lack of any discernible distributional effects associated with changes in returns to attributes suggests that this may be true. Another possible explanation, closely related to this last point, is that there may have been changes in returns to unobservable characteristics, which are orthogonal to the set of observable attributes. An alternative explanation for the (unexplained) changes in the top half of the income distribution, may be related to factors driving changes in unearned income.

V: Implications for Changes in Income Inequality Between 1983-86 and 1995-98

We now turn to the third feature of the analytical framework used here, which is that the various counterfactual distributions of income described above can be used to estimate the contribution of the each of the factors to level of, and change in, various summary measures of income inequality over the period. This provides a way to *quantify* the qualitative changes in the distribution due to each of the sets

[20] That is, the coefficient from the regression of the explained density change on the total density change is 0.65.

of factors considered as depicted visually in the various counterfactual distribution figures above.

Table 4 presents a summary of this analysis. In panel A we present, for the actual and each of the counterfactual distributions described above, the 10th, 25th, 50th (median), 75th, and 90th percentiles of the log-income distribution. These percentiles provide a quantitative summary of the various distributions, and comparing percentiles in adjacent columns of the table provides an indication of the estimated marginal effect of the various factors on the distribution observed in figures 5 – 9 above. For example, comparing the percentiles of the counterfactual distribution which allows for changing household structure (column 2) with the actual 1983-86 distribution (column 1) confirms that the estimated effect of changing household structure is to lower the distribution of household income: by about 5 percent (i.e. from $15,183 to $14,451) at the 10th percentile, by 6 percent at the 25th percentile, 4 percent at the median, 2 percent at the 75th percentile, and only 1 percent at the 90th percentile of the distribution. Similarly, and as expected, comparing columns 2 and 3 shows that changes in National Superannuation rates only affect the lower end of the distribution: raising income at the 10th percentile by 3 percent, and lowering income at the 25th percentile by 1 percent. In contrast, changing household socio-demographic attributes (column 4) is predicted to raise the distribution of income, especially at the higher percentiles: e.g., by 4 percent at the median, 6 percent at the 75th percentile, and 7 percent at the 90th percentile of the distribution. The effect of changing employment outcomes (column 5) is estimated to lower household incomes by 3-6 percent at each of the five percentiles, while changing returns to attributes (column 6) tends to raise incomes in the lower half of the distribution by 1-2 percent. Finally, comparing the final column with column 6, shows the unexplained changes: broadly speaking, except for the 90th percentile where there is a large (8 percent) unexplained increase in income, the combined factors we examine provide a reasonable account of the observed changes in the distribution.

In panel B of table 4, for each of the distributions, we present five summary measures of inequality commonly used in the literature. These are the Gini coefficient (Gini), the standard deviation of log-income (SD), the interquartile range of log-income (IQR), the difference between the 90th and 50th percentiles of the distribution (90-50), and the difference between the 50th and 10th percentiles of the

distribution (50-10). Comparing the entries in adjacent columns provides a similar interpretation as for the percentiles in panel A. Panel C summarises these changes in the measures of inequality. The first column presents the actual change in inequality between 1983-86 and 1995-98, together with the percentage change relative to 1983-86 in parentheses. The remaining columns contain the marginal contribution to the total change of each factor, together with the relative contribution in parentheses (and the final column contains the "unexplained" change).

There are several interesting and important points to note from panels B and C about the changing income inequality over this period. First, the increase in (actual) inequality varies according to the measure of inequality (column 1 of panel C): the Gini and IQR both increased by 15 percent over the period, the standard deviation of log income (SD) increased by 9 percent, while the 90-50 difference increased 26 percent, and there was actually a small (4 percent) drop in the 50-10 difference.[21] One interpretation of these differences is that the increase in dispersion in incomes was not even across the distribution.

Second, these factors account for a significant fraction of the increase in inequality over the period. Specifically, the fraction of the observed increase in inequality accounted for by the sets of factors we consider varies from between about 40 percent in the case of the Gini coefficient and the 90-50 difference, and 65 percent in the case of the interquartile range.

Third, changes in household structure and sociodemographic attributes tend to provide the largest marginal contributions to the change in inequality, however the actual contribution varies substantially depending on the specific factor and measure of inequality. Household structure changes explain between 14 percent (SD) and one-third (IQR) of the total increase in inequality over the period, while socio-demographic changes explain between 12 percent (90-50) and nearly one-half (SD) of the increase. In contrast, changes in National Superannuation, employment, and returns, each explain less than one-sixth

[21] Note, because of the relatively small change in the 50-10 difference, the factor-specific contributions in panel C will be quite sensitive to sampling variation. For this reason, we largely ignore changes in the 50-10 difference in the discussion, although the table presents the results for this measure. In this case we believe a more robust way to understand the causes of changes in the lower half of the distribution is to examine the relevant percentiles in panel A.

of the various increases. In particular, the results in table 5 support the earlier graphical finding that changes in the returns to household attributes had very little systematic effect on income inequality over this period. Furthermore, although changes in National Superannuation rates had a marked effect on the distribution of income, the effect was sufficiently localised that it did not translate into a substantial contribution to particular measures of inequality. However, employment changes do provide the single largest contribution to the increase in the top half of the distribution, as measured by the 90-50 difference.

The final row in panel C contains the Kullback-Leibler J-statistic, which provides an index of how dissimilar two densities are.[22] This index provides a measure that is closer in spirit to the density differences shown in figure 3b. The first column shows how dissimilar the 1983-86 and 1995-98 distributions are - an index of 3.37. In contrast to most of the inequality measures, the J statistic assigns an important role to National Superannuation rate changes (33.4%). This accords with the visual impact of the spikes in Figure 3b. Changes in household types, attributes, and employment each contribute between 7 and 11 percent of the change. The estimated impact of changing returns to attributes is to change the employment-adjusted density to be more like the 1983-86 distribution than like the 1995-98 distribution. Around half of the change is left unexplained.

While each of the summary statistics conveys some relevant information about the changing distribution of household incomes, there are clearly drawbacks in relying on any one measure. The graphical summaries derived from semiparametric methods provide much richer information on the nature of changes.

VI: Sub-period Changes in the Distribution of Household Incomes

The analysis described above has concentrated on decomposing the change in household incomes between the 1983-86 and 1995-98 periods. However, as we observed in section II, most of the changes

[22] The J-statistic associated with two distributions denoted "0" and "1" is calculated as

$$J_y^{01} = \int_0^\infty \{f_0(y) - f_1(y)\} \ln\{f_0(y)/f_1(y)\} dy .$$

appear to have occurred during the late 1980s. In particular, the dramatic increase in Gini coefficients and the shift in the distribution of household types were concentrated in this period. In addition, employment numbers fell strongly during this period, and recovered somewhat during the 1990s. To examine the effects of changes in the sets of factors of interest above over the 1980s and 1990s, in this section we repeat the analysis described above but applied to the two subperiods between 1983-86 and 1989-92, and 1989-92 and 1995-98. For each subperiod, we have decomposed changes relative to the initial period – i.e. 1983-86 acts as the base period for the analysis of changes between 1983-86 and 1989-92, and 1989-92 acts as the base period for the analysis of changes between 1989-92 and 1995-98.

We begin by comparing the actual household income distributions in 1983-86 and 1989-92, and in 1989-92 and 1995-98, shown in figures 12a and 12b respectively. These figures support the view that most of the change in the income distribution appears to have occurred during the 1980s. Figure 12a shows that the hollowing out of the middle-income range, and the redistribution to the low and high ends of the distribution, occurred mainly between 1983-86 and 1989-92. In contrast, figure 12b shows there is relatively little change in this respect between 1989-92 and 1995-98. However, the shifts in the "National Superannuation" spikes in the distribution appear to have occurred in the 1990s.

We next highlight the analytical results for each sub-period. For the first sub-period between 1983-86 and 1989-92, figure 13a describes the 1989-92 counterfactual distribution which accounts for changes in all of the various factors together with the actual 1989-92 distribution, while figure 13b describes the explained change in the distribution together with the total change over the period.[23] These figures suggest that the various sets of factors provide quite a reasonable account of the actual changes in income distribution over the late 1980s. In particular, the various factors provide a good description of the downward shifts from the middle to the low end of the income distribution. However, the shift upward is again less well explained by the factors we examine.

Figure 12b suggests that the principal change in the distribution of income over the second sub-

[23] In order to save on space, we present a summary of the sub-period analyses. The entire set of figures corresponding to the various counterfactual stages is available on request from the authors.

period, between 1989-92 and 1995-98, may have been due to changes in National Superannuation rates. In addition to the National Superannuation related changes, there is also a discernible increase in the density at the high end of the distribution: as before this change remains largely unexplained by the factors we have examined. In order to concentrate on the effects of National Superannuation, in figure 14a we show the counterfactual distribution which allows for changes in the distribution of household types and changes in the statutory rates of National Superannuation, together with the actual 1995-98 distribution of household income, while figure 14b shows the marginal effect of changing National Superannuation together with the total change in the distribution over this period. These figures support the view that the main changes in the income distribution over this period were due to National Superannuation changes.

We now consider the implications of the analysis for the various summary measures of inequality over each of the two sub-periods. Table 5 summarises the decomposition in the alternative measures of inequality attributable to the various sets of factors for the two subperiods. As suggested by figures 12a and 12b, the change in inequality is largely concentrated in the first sub-period. The results for the first sub-period tell a remarkably consistent story to that for the full period shown in table 4. Changes in household structure and attributes again provide the largest contributions to the change in inequality, with the exception of employment changes which explain over 20 percent of the increase in the 90-50 difference during the 1980s. Also, the results in this table confirm that the sets of factors account for a larger fraction of the increase in inequality over this period than over the entire sample period: the fraction of inequality "explained" by these factors ranges from 60 percent in the case of the Gini up to over 80 percent in the case of the interquartile range. The results for the second sub-period confirm that there was relatively little change in inequality over this period: less than a 5 percent change in inequality using any measure.

VII: Changes in the Household Type Distributions of Income

In this section, we summarise the results of the analysis of changes in the sub-distributions of income for each of the six household types described above. This analysis also provides a decomposition of how various changes in the overall distribution of household income are attributable to changes in the distribution of income within the various household types, as well as providing an alternative to equivalised income for the whole population of households.

Table 6 summarises the decomposition in the alternative measures of inequality attributable to the various sets of factors for each household type over the full period 1983-86 to 1995-98. The first column presents, for each household type, the actual change in each measure of inequality over the period together with the relative change in parentheses, while subsequent columns present the relative marginal contribution of the respective factors. The results here provide further evidence on the complexity of the changing distribution of household incomes over this period.

First, there were quite diverse changes in income inequality across the various household types. With the exception of the "single adult with children" households, there were strong increases in income inequality for the under-60s household types. For example, the Gini coefficient increased by about 20 percent for each of these three groups. In contrast, the level of inequality among the "single adult with children" households actually fell: by 14 percent for the Gini coefficient by, and between 5 percent and one-third, using other measures.

Second, although changing employment outcomes had a relatively modest effect on the changes in inequality at the aggregate household level, this factor contributed the strongest effect to changes in inequality within each household type. For example, employment changes account for 43 percent, 30 percent and 16 percent of the increase in the Gini for the single adult, multiple adult, and multiple adults with children households respectively, and two-thirds of the decrease among the single adult with children households. In addition, changes in returns also contributed significantly to the changes in inequality within household groups.

The changes for the two over-60s household types are less consistent across measures. For

example, the Gini coefficient and standard deviation of log income measures of inequality both increased (by 1 to 11 percent), while the interquartile range and 50-10 difference measures decreased (by 3 to 11 percent). As expected, National Superannuation changes contribute significantly to these changes. However, there also seem to be other factors at work, particularly in the case of the multiple adult households. For example, although the IQR fell by 9 percent, the effects of National Superannuation changes were to actually increase this measure. Although we haven't explored other factors for these groups, given the increasing age of eligibility, employment effects plausibly contribute to the changes for this group.

VIII: Concluding Discussion

In this paper we have presented a semiparametric analysis of the effects of various quantifiable factors on changes in the distribution of household incomes in New Zealand over the period between 1983 and 1998. The analysis facilitates a visual appreciation of, first, the distribution of income within a period, second, changes in the distribution between periods, and third, the contributions of the various factors to the level and change in the distribution. In addition, the analysis enables a decomposition of changes in alternative summary measures of inequality into effects attributable to the various factors.

As has been documented in other research, we find that the increase in income inequality was concentrated during the late 1980s. However, our analysis shows that the changes in the distribution of income involved a complex set of factors which are difficult to summarise using a single measure of inequality. Examining income inequality across all households, we find that the main factors which contributed to the change in inequality were changes in family and household structure (primarily a pronounced drop in the fraction of two parent households and a rise in the fraction of sole parent households), and changes in the socio-demographic attributes of households. These factors each explain one-sixth of the total increase in the Gini coefficient over the period, and up to one-third and one-half (respectively) of other measures of inequality. In addition to these factors, changes in National Superannuation rates had prominent effects on the distribution of income; however these effects were

localised and had relatively little impact on the summary measures of inequality that we examine. Somewhat surprisingly, our results show that changes in the employment outcomes of households had a more modest impact on income inequality. However, within household types, we find that employment changes do have a large effect on the observed change in income inequality. Finally, we find little evidence of any systematic effects of changes in the economic returns to socio-demographic attributes on the distribution of household income and inequality.

There are several important caveats to the analysis and results. First, as the analysis of the effects of the different sets of factors is sequential in nature, the marginal effects of each factor will depend on the sequential ordering of the analysis (unless changes across factors are orthogonal). The (primary) ordering used in the analysis is chosen to reflect (we hope) a progressively dependent ordering of the various factors. Also, as a check of the robustness of the results, we plan to reverse the sequential ordering and repeat the analysis.

Second, although the results could be viewed as the effects of changes in exogenous factors on the distribution of income, this interpretation would be somewhat naïve. In particular, changes in household structure, attributes and employment may be, at least partially, endogenous to changes in the income distribution, and may also reflect the effects of some other underlying factors. Rather, our view is that the analysis provides an account of the possible effects of several potentially important sets of factors which have an influence on the distribution of income.

Third, the analysis focuses on gross household income, which neglects several potentially important effects and changes on the welfare of households and individuals over the period. Among the neglected issues are the equivalisation metric to facilitate welfare comparisons of different household structures. Foremost among the changes is the impact of tax reform and fiscal incidence changes over the period, which may have dramatically and differentially affected the level of disposable income available to households at different points in the income distribution. The impacts of these issues remain open to future analysis.

References

Coleman, Andrew, and Paul Hansen (1996), "", Manuscript, New Zealand Treasury.

Daly, Mary C., and Robert G. Valletta (2000), "Inequality and Poverty in the United States: The Effects of Changing Family Behavior and Rising Wage Dispersion", Manuscript, Federal Reserve Bank of San Francisco.

Davey, Judith (1998), *Tracking Social Change in New Zealand*, Institute of Policy Studies, Victoria University, Wellington.

Dixon, Sylvia (1996), "Labour Force Participation Over the Last Ten Years", *Labour Market Bulletin*, pp. 71-88.

Dixon, Sylvia (1998), "Growth in the Dispersion of Earnings: 1984-97", *Labour Market Bulletin*, pp. 71-107.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux (1996), "Labor Market Institutions and the Distribution of Wages, 1973—1992: A Semiparametric Approach", *Econometrica*, Vol. 64, pp. 1001-1044.

Easton, Brian (1996), "Income Distribution", in Brian Silverstone, Alan Bollard and Ralph Lattimore (eds), *A Study of Economic Reform: The Case of New Zealand*, North-Holland.

Frame, David (1999), "Labour Force Participation and Superannuation", Manuscript, New Zealand Treasury.

Maani, S. (1999), *Private and Public Returns to Investments in Secondary and Higher Education in New Zealand over time: 1981-1996,* Treasury Working Paper #99/ [available from http:\\www.treasury.govt.nz]

Martin, Barry (1998), *Incomes of Individuals and Families in New Zealand, 1951 to 1996.* Unpublished thesis, University of Waikato.

O'Dea, Des (2000), "The Changes in New Zealand's Income Distribution", New Zealand Treasury Working Paper, 00/13 [available from http:\\www.treasury.govt.nz].

Podder and Chatterjee (1998), "Sharing the National Cake in Post Reform New Zealand: Income Inequality Trends in Terms of Income Sources", presented to the NZAE Conference, Wellington.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.

Silverstone, Brian, Alan Bollard and Ralph Lattimore (eds) (1996), *A Study of Economic Reform: The Case of New Zealand*, North-Holland.

Statistics New Zealand (1998), *New Zealand Now - Families and Households,* Wellington.

Statistics New Zealand (1999), *New Zealand Now - Incomes*, Wellington.

Appendix: Weighted Kernel Density Estimation and Counterfactual Distributions

In this appendix we first provide a brief outline of the technique used to obtain the Kernel density estimates of the income distributions in the paper, and second provide details of the construction of the counterfactual distributions for changes in the various sets of factors. The description of the analysis is analogous to that presented in DiNardo, Fortin and Lemieux (1996); while Silverman (1986) provides a detailed accounted of kernel density estimation.

A: Weighted Kernel Density Estimation

Let y represent the natural logarithm of household income, and let $f_t(y)$ be the probability density associated with log household income y in time period t. Kernel density methods for estimating $f_t(y)$ compute a weighted average of the observations *near* to y, where the weighted averaging is determined by a Kernel function K(.), and *near* is defined by a bandwidth (or *window-width*) h. In particular, given a random sample of N observations in period t (y_{t1}, ..., y_{tN}), with sampling weights (θ_{t1}, ..., θ_{tN}) ($\sum_{i=1}^{N} \theta_{ti}=1$), the weighted Kernel density estimate of $f_t(y)$ is

$$\hat{f}_t(y)=\sum_{i=1}^{N} \frac{\theta_{ti}}{h} \cdot K\left(\frac{y - y_{ti}}{h}\right) \tag{A1}$$

where h is the bandwidth, and K(.) is the kernel function. Throughout the analysis, we use a bandwidth h=0.05, and the Epanechnikov kernel:

$$K(z) = \begin{cases} 3/4(1 - z^2/5)/\sqrt{5} & if\ |z| < \sqrt{5} \\ 0 & otherwise \end{cases}.$$

The choice of bandwidth tends to be the more important issue in kernel density estimation. Our choice of h=0.05 is somewhat narrower than suggested as an "optimal" bandwidth, which assumes that the underlying distribution is smooth. If this assumption is not valid, then the estimated optimal bandwidth will tend to "over-smooth" the distribution and disguise localised peaks and troughs. It is largely this observation that motivates the use of a narrower bandwidth. In addition, it is simpler for the reader to

"smooth" (by-eye) an estimated distribution that is based on a too-narrow bandwidth, than it is to "unsmooth" a distribution that is based on a bandwidth that is too-wide. As a partial check on the robustness of the results we have re-estimated some of the distributions using wider bandwidths and found the results are qualitatively unchanged. All of the estimated distributions are evaluated at 250 equi-spaced log income values between 7.8 and 12.5.

B: Counterfactual Distribution Construction

We now explain the sequence of counterfactuals constructed to analyse changes in the distribution of household income between period "0" and "1" (primarily between 1983-86 and 1995-98, but alternatively between 1983-86 and 1989-92, and between 1989-92 and 1995-98). This involves a combination of conditionally translating (in the case of National Superannuation and economic returns adjustments), and reweighting (in the case of household types, attributes and employment outcomes) empirical distributions to take account of changes in the various sets of factors of interest.

Household Types

As described in the text, $f_t(y)$ can be expressed as the weighted average density across the J distinct household types:

$$f_t(y) = \sum_{j=1}^{J} w_{tj} \cdot f_{tj}(y) \qquad \text{(A2)}$$

where w_{tj} is the fraction of households of type j in period t ($w_{tj} = \sum_{i=1}^{N_{tj}} \theta_{ti}$, and N_{tj} is the number of households of type j in period t), and $f_{tj}(y)$ is the probability density at log income y of household type j in period t. The counterfactual distribution of income which allows the distribution of households to change from period 0 to period 1, but holds the distribution of household-type incomes constant between period 0 and 1 is simply

$$\hat{f}_0^H(y) = \sum_{j=1}^{J} w_{1j} \cdot \hat{f}_{0j}(y). \tag{A2a}$$

That is, $\hat{f}_0^H(y)$ is obtained by reweighting the period-0 household-type income densities by the period-1 household-type fractions, and the estimated change in the distribution explained by this counterfactual is

$$\Delta\hat{f}^H(y) = \hat{f}_0^H(y) - \hat{f}_0(y) = \sum_{j=1}^{J} (w_{1j} - w_{0j})\hat{f}_{0j}(y). \tag{A2b}$$

National Superannuation

Conditional on changes in the distribution of household-types, the effects of changes in the statutory rate(s) of National Superannuation over the period can be expressed in terms of translations of incomes as follows. Let Y_{ti} be the level of income of household i in year t (i.e. $y_{ti} = \log(Y_{ti})$), and let N_{ti} be the National Superannuation component of income for the household. Constructing a counterfactual distribution for changes in National Superannuation rates requires adjusting the actual NS received in period 0 (N_{0i}) to take account of the statutory changes between period 0 and period 1. We denote this change $\pi_i = NS_{1i}/NS_{0i}$, where NS_{0i} is the statutory rate of National Superannuation applicable to household i in period 0,[24] and NS_{1i} is the statutory rate for this household exactly 12 years later (6 years later in the 1983/86 – 1989/92 and 1989/92 – 1995/98 sub-period analyses). We then scale up (or down) the actual National Superannuation received in period 0 to account for the statutory-rate change, and adjust (log) total income for the household in period 0, $\hat{y}_{0i}^N = \log(Y_{0i} + (\pi_i - 1)N_{0i})$. For each household type j, we then obtain kernel density estimates of the distribution of $\hat{y}_{0i}^N$,

$$\hat{f}_{0j}^N(y) = \sum_{i=1}^{N_j} \frac{\theta_{0i}}{h} \cdot K\left(\frac{y - \hat{y}_{0i}^N}{h}\right),$$

and the distribution of income across all households,

[24] For households with one person aged over 60 we apply the statutory rate for single people, while for households with at least two people aged over 60 we apply the statutory rate for couples. If the household has no one aged over 60,

$$\hat{f}_0^{NH}(y) = \sum_{j=1}^{J} w_{1j} . \hat{f}_{0j}^{N}(y) . \tag{A3a}$$

The marginal effect of the change in the distribution explained by changes in National Superannuation rates is

$$\Delta\hat{f}^{N}(y) = \hat{f}_0^{NH}(y) - \hat{f}_0^{H}(y) = \sum_{j=1}^{J} w_{1j} . (\hat{f}_{0j}^{N}(y) - \hat{f}_{0j}(y)) . \tag{A3b}$$

Household Socio-demographic Attributes and Employment Outcomes

Conditional on changes in the distribution of household-types and statutory rates of National Superannuation, the effects of changes in the socio-demographic attributes and employment outcomes of households on the distribution of household income can be captured by reweighting the period 0 sample weights to take account of changes in such attributes and employment outcomes between period 0 and period 1.

To facilitate this, we can express the household-type density of income as the integral of the density of income conditional on a set of household demographic attributes, x, and employment outcomes, e:

$$f_{tj}(y) = \int_{(e,x)\in\Omega_{(e,x)}} f_{tj}(y;x,e)dF_{tj}(e,x) \tag{A4}$$

where $\Omega_{(e,x)}$ is the domain of the household employment outcomes and demographic attributes, and $F_{tj}(e,x)$ is the joint distribution of (e,x) of household type j in period t. In order to distinguish the effects of changes in the attributes and employment outcomes, we consider a sequential analysis that analyses the effects of, first, changes in demographics attributes and, second, changes in employment outcomes conditional on attributes. This approach is captured by rewriting (A4) as

$$f_{tj}(y) = \int_{x\in\Omega_x} \int_{e\in\Omega_{e|x}} f_{tj}(y;x,e).dF_{tj}(e \mid x)\, dF_{jt}(x) . \tag{A5}$$

For each household-type j, we consider the effects of changes in household demographic attributes, conditional on changes in the statutory National Superannuation rates. In particular, we construct a

---

we set π_{ji}=0.

counterfactual density allowing the distribution of attributes to be as observed in period 1, *but* the conditional distribution of employment outcomes *and* the density of incomes within each household type remained as in period 0:

$$f_{0j}^{XN}(y) = \int_{x\in\Omega_x}\int_{e\in\Omega_{e|x}} f_{0j}^{N}(y;x,e).dF_{0j}(e \mid x)\, dF_{1t}(x)$$

$$= \int_{x\in\Omega_x}\int_{e\in\Omega_{e|x}} f_{0j}^{N}(y;x,e).dF_{0j}(e \mid x).\psi_{xj}(x)\, dF_{0t}(x), \qquad \text{(A6)}$$

where $\psi_{xj}(x)=dF_{1j}(x)/dF_{0j}(x)$ is a "reweighting" function which rescales the period 0 density of attributes to obtain the prevailing period 1 density. In order to obtain an estimate of this reweighting function note that, by applying Bayes' rule, this function can be expressed as

$$\psi_{xj}(x) = \frac{P_j(t=1 \mid x)}{P_j(t=0 \mid x)} \cdot \frac{P_j(t=0)}{P_j(t=1)}$$

where $P_j(t=1|x)$ is the conditional probability that a household with attributes x is observed in period 1, and $P_j(t=1)$ is the unconditional probability that the household is observed in period 1.

In order to obtain estimates of the reweighting function, for each household type j, we first pool the period 0 and 1 households and estimate the probability that household i is observed in period 1, given attributes x, using a logit model for the binary dependent variable t.[25] We then use the estimates from this model to predict, for each household observed in period 0, the relative probability that it would be observed in period 1 versus period 0 (i.e. $\hat{P}_j(t=1 \mid x_{0i}) / \hat{P}_j(t=0 \mid x_{0i})$) and adjust this by $P_j(t=0)/P_j(t=1)$, to obtain the estimated "reweight" for this household, $\hat{\psi}_{xj}(x_{0i})$. The counterfactual density for household type j that takes account of changes in attributes, is then estimated by

$$\hat{f}_{0j}^{XN}(y) = \sum_{i=1}^{N_j} \frac{\hat{\psi}_{xj}(x_{0i}).\theta_{0i}}{h} . K\left(\frac{y - \hat{y}_{0i}^{N}}{h}\right),$$

and the counterfactual distribution of income across all households is estimated by

[25] The logit specifications and model results are presented in column (1) of appendix tables A1(a) – A1(d) for each

$$\hat{f}_0^{XNH}(y) = \sum_{j=1}^{J} w_{1j} . \hat{f}_{0j}^{XN}(y) \quad . \tag{A7a}$$

Intuitively, this reweighting scheme puts more weight on households with attributes that are more likely to occur in period 1 and less weight on households with attributes that are less likely to occur in period 1. Comparing $f_0^{XNH}(y)$ and $f_0^{NH}(y)$ provides a way to estimate the marginal effect of the change in the distribution that is explained by changes in socio-demographic attributes ($\Delta f^X(y)$). That is,

$$\Delta \hat{f}^X(y) = \hat{f}_0^{XNH}(y) - \hat{f}_0^{NH}(y) = \sum_{j=1}^{J} w_{1j} . (\hat{f}_{0j}^{XN}(y) - \hat{f}_{0j}^{N}(y)) \quad . \tag{A7b}$$

Returning now to the effects of changes in employment outcomes conditional on changes in attributes. In this case, for each household type, we construct a counterfactual density that allows the distribution of employment outcomes to be as in period 1, *but* the density of incomes conditional on employment to be as in period 0. That is,

$$f_{0j}^{EXN}(y) = \int_{x \in \Omega_x} \int_{e \in \Omega_{elx}} f_{0j}^{N}(y; x, e) . dF_{1j}(e \mid x)\, dF_{1j}(x)$$

$$= \int_{x \in \Omega_x} \int_{e \in \Omega_{elx}} f_{0j}^{N}(y; x, e) . \psi_{elx,j}(e, x) dF_{0j}(e \mid x) . \psi_{xj}(x)\, dF_{0j}(x) , \tag{A8}$$

where $\psi_{elx,j}(e,x)=dF_{j1}(elx)/dF_{j0}(elx)$ is a "reweighting" function which rescales the period-0 density of employment outcomes conditional on attributes to obtain the prevailing period-1 density. For this purpose, we consider a discrete set of M_j employment outcomes, defined according to the number of fulltime equivalent workers in the household and the type of household, j, and define $e_m=1$ if the household has employment outcome m and $e_m=0$ otherwise (m=0, …, M_j). In this case,

$$\psi_{elx,j}(e,x) = \sum_{m=0}^{M_j} e_m \cdot \frac{P_{1j}(e_m = 1 \mid x)}{P_{0j}(e_m = 1 \mid x)}$$

where $P_{tj}(e_m=1|x)$ is the probability of employment outcome m in period t, given household demographic attributes x.

---

household type.

To estimate the reweighting function $\psi_{elx,j}(e,x)$, we first estimate, separately for each period, either logit (if $M_j=1$) or ordered logit (if $M_j>1$) models for the employment outcome conditional on attributes.[26] For each household observed in period 0, we then use these models to predict the relative probability of employment outcome e_m in period-1 versus period-0 (i.e. $\hat{P}_{1j}(e_m=1\mid x_{0i})/\hat{P}_{0j}(e_m=1\mid x_{oi})$) to obtain the estimated reweight for this household, $\hat{\psi}_{elx,j}(e,x_{0i})$. Given this estimated reweight, the counterfactual density for household type j that takes account of changes in employment outcomes, is estimated by

$$\hat{f}_{oj}^{EXN}(y)=\sum_{i=1}^{N_{0j}}\frac{\hat{\psi}_{elx,j}(e,x_{0i}).\hat{\psi}_{xj}(x_{0i}).\theta_{0i}}{h}.K\left(\frac{y-\hat{y}_{0i}^{N}}{h}\right).$$

The counterfactual distribution of income across all households is again obtained by taking the weighted average across household types:

$$\hat{f}_0^{EXNH}(y)=\sum_{j=1}^{J}w_{1j}.\hat{f}_{0j}^{EXN}(y), \tag{A9a}$$

and the marginal effect of the change in the distribution explained by changes in employment outcomes is

$$\Delta\hat{f}^{E}(y)=\hat{f}_0^{EXNH}(y)-\hat{f}_0^{XNH}(y)=\sum_{j=1}^{J}w_{1j}.(\hat{f}_{0j}^{EXN}(y)-\hat{f}_{0j}^{XN}(y)). \tag{A9b}$$

Economic Returns to Attributes

The final explanatory factor that we analyse is changes in the economic "returns" to attributes, conditional on changes in household type, National Superannuation, demographic attributes and employment outcomes. For this exercise, we construct a counterfactual density allowing the income returns to observed household socio-demographic attributes to be as in period-1, by adjusting each household's period-0 income by the predicted change given their attributes.

In particular, for each household-type j, we first estimate regressions of log-income on socio-demographic attributes separately for each period:

[26] The employment outcomes are described in table 2 for each household type. Summaries of the model results are

$$\hat{y}_{tji}^{N} = X_{tji}^{'}\beta_{tj} + \varepsilon_{tji}.$$[27]

We then compute the predicted change in returns, $\Delta\hat{y}_{ji} = X_{0ji}^{'}(\hat{\beta}_{1j} - \hat{\beta}_{0j})$, and log household income adjusted for this change, $\hat{y}_{0ji}^{R} = \hat{y}_{0ji}^{N} + \Delta\hat{y}_{ji}$, and then obtain the counterfactual density for household-type j that takes account of this change in returns:

$$\hat{f}_{oj}^{REXN}(y) = \sum_{i=1}^{N_j} \frac{\hat{\psi}_{elx,j}(e, x_{0i}).\hat{\psi}_{xj}(x_{0i}).\theta_{0i}}{h}.K\left(\frac{y - \hat{y}_{0i}^{R}}{h}\right),$$

and the counterfactual distribution of income across all households is again obtained by taking the weighted average across household types:

$$\hat{f}_{0}^{REXNH}(y) = \sum_{j=1}^{J} w_{1j}.\hat{f}_{0j}^{REXN}(y), \tag{A10a}$$

and the marginal effect of the change in the distribution explained by changes in economic returns is

$$\Delta\hat{f}^{R}(y) = \hat{f}_{0}^{REXNH}(y) - \hat{f}_{0}^{EXNH}(y) = \sum_{j=1}^{J} w_{1j}.(\hat{f}_{0j}^{REXN}(y) - \hat{f}_{0j}^{EXN}(y)). \tag{A10b}$$

On the basis of these sequentially constructed counterfactual densities, the total change in the density of log household income between period-0 and period-1 can be decomposed into "explained" and "unexplained" components, as follows:

$$(\hat{f}_{1}(y) - \hat{f}_{0}(y)) = (\hat{f}_{0}^{REXNH}(y) - \hat{f}_{0}(y)) + (\hat{f}_{1}(y) - \hat{f}_{0}^{REXNH}(y)) \tag{A11}$$

where $\hat{f}_{0}^{REXNH}(y) - \hat{f}_{0}(y)$ represents the total change in the distribution explained by the sets of factors we have examined here, and $\hat{f}_{1}(y) - \hat{f}_{0}^{REXNH}(y)$ represents the change in the overall income distribution which remains unexplained.

---

presented in columns (2) and (3) of appendix tables A1(a) – A1(d) for each household type.

[27] We estimate these models using Quantile (median) regression rather than least squares methods in order to minimise the effects of outliers on the estimated shifts in the returns. The results of these models are presented in columns (4) and (5) of appendix tables A1(a) – A1(d) for each household type.

Table 1: Sample Characteristics

	Single Adult Over 60			Multiple Adults Over 60		
	1983-86	**1989-92**	**1995-98**	**1983-86**	**1989-92**	**1995-98**
Mean Household Income	20,389	19,688	21,547	39,685	41,378	38,355
	(527)	(649)	(690)	(755)	(918)	(909)
Median Income	14,638	14,089	15,342	30,116	30,495	28,143
Gini Coefficient	0.291	0.269	0.303	0.302	0.324	0.328
No. Persons	1	1	1	2.04	2.04	2.04
				(.005)	(.006)	(.006)
Age of Adults	71.6	72.9	73.8	66.8	66.9	67.3
	(.22)	(.22)	(.24)	(.18)	(.19)	(.20)
Fraction of Adults:						
Female	0.71	0.74	0.74	0.51	0.51	0.51
	(.01)	(.01)	(.01)	(.002)	(.002)	(.002)
Married	0.02	0.02	0.03	0.92	0.92	0.93
	(.004)	(.004)	(.01)	(.01)	(.01)	(.01)
Maori	0.01	0.03	0.04	0.02	0.03	0.04
	(.003)	(.005)	(.01)	(.003)	(.004)	(.005)
Employed Fulltime	0.07	0.03	0.04	0.13	0.10	0.14
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Employed Parttime	0.03	0.03	0.02	0.04	0.06	0.09
	(.005)	(.01)	(.004)	(.004)	(.005)	(.01)
Fraction of Households	0.106	0.123	0.115	0.133	0.146	0.132
Sample size	1,117	1,174	1,059	1,409	1,326	1,144

Notes: Standard errors are in parentheses. All means are weighted by the HES sample weights. Incomes are measured in constant (1999) dollars.

Table 1: Sample Characteristics
(Continued)

	Single Adult Under 60			Multiple Adults Under 60		
	1983-86	**1989-92**	**1995-98**	**1983-86**	**1989-92**	**1995-98**
Mean Household	32,354	35,647	35,291	67,125	69,602	70,552
Income	(1,014)	(2,260)	(1,073)	(731)	(1,007)	(1,051)
Median Income	28,414	28,271	28,772	61,636	61,791	61,384
Gini Coefficient	0.328	0.375	0.390	0.272	0.311	0.326
No. Persons	1	1	1	2.47	2.47	2.47
				(.01)	(.02)	(.02)
Age of Adults	39.8	41.8	42.6	36.3	37.1	39.0
	(.44)	(.40)	(.39)	(.22)	(.23)	(.23)
Fraction of Persons:						
Aged 15-18	0.01	0.01	0.01	0.06	0.06	0.04
	(.003)	(.003)	(.003)	(.003)	(.003)	(.002)
Fraction of Adults[a]:						
Female	0.44	0.47	0.45	0.49	0.49	0.49
	(.02)	(.02)	(.02)	(.004)	(.004)	(.004)
Married	0.01	0.03	0.03	0.58	0.56	0.68
	(.004)	(.01)	(.01)	(.01)	(.01)	(.01)
Maori	0.04	0.06	0.08	0.04	0.06	0.07
	(.01)	(.01)	(.01)	(.003)	(.004)	(.004)
School	0.24	0.25	0.28	0.28	0.34	0.33
Qualifications	(.02)	(.02)	(.02)	(.01)	(.01)	(.07)
Vocational	0.28	0.29	0.27	0.24	0.26	0.27
Qualifications	(.02)	(.02)	(.02)	(.01)	(.01)	(.01)
University	0.12	0.12	0.19	0.08	0.11	0.14
Qualifications	(.01)	(.01)	(.01)	(.004)	(.01)	(.01)
Employed	0.79	0.65	0.72	0.76	0.67	0.68
Fulltime	(.01)	(.02)	(.02)	(.01)	(.01)	(.01)
Employed	0.03	0.06	0.09	0.05	0.11	0.13
Parttime	(.01)	(.01)	(.01)	(.003)	(.004)	(.004)
Fraction of Households	0.077	0.088	0.091	0.265	0.254	0.277
Sample size	814	829	794	2,802	2,340	2,383

Notes: Standard errors are in parentheses. All means are weighted by the HES sample weights. Incomes are measured in constant (1999) dollars.

[a] Adults are defined as those persons aged at least 18 or aged 15-18 with positive wage and salary or benefit income.

Table 1: Sample Characteristics
(Continued)

	Single Adult With Children			Multiple Adults With Children		
	1983-86	1989-92	1995-98	1983-86	1989-92	1995-98
Mean Household Income	24,771	24,121	26,011	55,180	60,289	62,369
	(916)	(538)	(1,085)	(622)	(809)	(1,021)
Median Income	19,575	21,821	22,302	49,905	52,235	50,837
Gini Coefficient	0.297	0.283	0.255	0.285	0.321	0.345
No. Persons	2.87	2.80	2.92	4.41	4.30	4.32
	(.05)	(.04)	(.04)	(.02)	(.02)	(.02)
Age of Adults	34.9	34.4	35.0	34.4	35.3	36.1
	(.44)	(.36)	(.34)	(.11)	(.13)	(.14)
Fraction of Persons:						
Aged 0-18	0.62	0.61	0.62	0.50	0.48	0.49
	(.01)	(.004)	(.005)	(.002)	(.002)	(.002)
Fraction of Adults[a]:						
Female	0.89	0.83	0.86	0.50	0.51	0.52
	(.01)	(.02)	(.01)	(.002)	(.002)	(.002)
Married	0.02	0.04	0.05	0.85	0.80	0.85
	(.01)	(.01)	(.01)	(.005)	(.01)	(.01)
Maori	0.13	0.24	0.24	0.10	0.12	0.12
	(.02)	(.02)	(.02)	(.004)	(.01)	(.01)
School Qualifications	0.25	0.28	0.30	0.23	0.32	0.33
	(.02)	(.02)	(.02)	(.005)	(.01)	(.01)
Vocational Qualifications	0.15	0.16	0.21	0.24	0.26	0.26
	(.02)	(.01)	(.02)	(.005)	(.01)	(.01)
University Qualifications	0.05	0.04	0.08	0.07	0.09	0.15
	(.01)	(.01)	(.01)	(.003)	(.004)	(.01)
Employed Fulltime	0.30	0.19	0.19	0.63	0.53	0.53
	(.02)	(.02)	(.02)	(.004)	(.01)	(.01)
Employed Parttime	0.11	0.13	0.20	0.11	0.16	0.19
	(.01)	(.01)	(.02)	(.003)	(.004)	(.005)
Fraction of Households	0.040	0.059	0.067	0.379	0.329	0.318
Sample size	428	605	585	4,008	3,022	2,733

Notes: Standard errors are in parentheses. All means are weighted by the HES sample weights. Incomes are measured in constant (1999) dollars.

[a] Adults are defined as those persons aged at least 18 or aged 15-18 with positive wage and salary or benefit income.

Table 2: Counterfactual Income Measures and Conditioning Weights

Distribution	Income Measure	Conditioning Weight
1. Base Period (t=0) Actual Distribution	y_{0ji}	θ_{0ji}
2. Changes in National Superannuation[(a)]	$\hat{y}^{N}_{0ji} = \log(Y_{0ji} + (\pi_i - 1)N_{0ji})$	θ_{0ji}
3. Changes in Socio-demographic Attributes[(b)]	$\hat{y}^{N}_{0ji}$	$\hat{\theta}^{X}_{0ji} = \hat{\varphi}_{xj}(x_{0ji}).\theta_{0ji}$
4. Changes in Employment[(c)]	$\hat{y}^{N}_{0ji}$	$\hat{\theta}^{EX}_{0ji} = \hat{\varphi}_{e\|x,j}(e, x_{0ji}).\hat{\theta}^{X}_{0ji}$
5. Changes in Economic Returns[(d)]	$\hat{y}^{RN}_{0ji} = \hat{y}^{N}_{0ji} + X'_{0ji}(\hat{\beta}_{1j} - \hat{\beta}_{0j})$	$\hat{\theta}^{EX}_{0ji}$
6. Final Period (t=1) Actual Distribution	y_{1ji}	θ_{1ji}

Notes: The entries describe the income measures and weights used in constructing the kernel density estimates of the various income distributions for household type j, where Y_{0ji} is the gross income of household i in household-type j in period 0, $y_{0ji} = \log(Y_{0ji})$, and θ_{0ji} is the sampling weight for that household. To construct the distribution of income across all households, the household type distributions are weighted by w_{0j} (period 0 actual distribution, row 1), or w_{1j} (counterfactual distributions, rows 2-5; and period 1 actual distribution, row 6), where $w_{tj} = \sum_{i=1}^{N_{tj}} \theta_{tji}$ is the sample fraction of households in type j in period t.

[(a)] $\pi_i = NS_{1i}/NS_{0i}$, where NS_{ti} is the rate of national superannuation applicable to household i in period t; and N_{0ji} is the actual national superannuation income of household i in period 0 – see text for details.

[(b)] $\hat{\varphi}_{xj}(x_{0ji})$ is the reweighting to adjust for changes in socio-demographic attributes between period 0 and 1 – see appendix for details.

[(c)] $\hat{\varphi}_{e|x,j}(e, x_{0ji})$ is the reweighting to adjust for changes in employment outcomes between period 0 and 1 – see appendix for details.

[(d)] $\hat{\beta}_{tj}$ are the estimated coefficients from median regressions of $\hat{y}^{N}_{0ji}$ regressed on the socio-demographic attributes X_{tji} – see appendix for details.

Table 3: Employment Outcomes

Employment Outcome	Single Adult Under 60 (1)	Multiple Adults Under 60 (2)	Single Adult With Kids (3)	Multiple Adults With Kids (4)
A: Employment Outcomes Descriptions (Numbers of FTE Workers)				
0	< 1	< 1	< 1	< 1
1	1	$\geq 1, < 2$	1	$\geq 1, < 2$
2	---	≥ 2	---	≥ 2
B: Frequency Distributions				
1983-86				
0	0.208	0.047	0.699	0.045
1	0.792	0.264	0.301	0.541
2	---	0.690	---	0.414
1989-92				
0	0.347	0.096	0.815	0.122
1	0.653	0.307	0.185	0.569
2	---	0.597	---	0.309
1995-98				
0	0.284	0.095	0.814	0.124
1	0.716	0.295	0.186	0.552
2	---	0.610	---	0.324

Notes: Non-workers, parttime workers and fulltime workers are classified as 0, 0.5 and 1 fulltime equivalent (FTE) worker respectively, and the employment outcome groupings for each household type are defined as shown in panel A.

Table 4: Implications for Measures of Income Inequality

	Actual 1983-86 Distribution	Counterfactual Distribution Allowing Changes in					Actual 1995-98 Distribution
		Household Structure	+National Superannuation	+Household Attributes	+Employment Outcomes	+Economic Returns	
				A: Percentiles of the Income Distribution			
10th	15,183	14,451	14,964	14,853	14,450	14,570	14,465
25th	25,275	23,649	23,407	23,726	22,837	23,288	22,607
50th	43,431	41,782	41,675	43,563	41,083	41,488	39,805
75th	64,966	63,683	63,613	67,561	65,301	65,538	66,874
90th	88,807	87,936	87,737	93,778	91,232	91,582	98,249
				B: Summary Measures of Income Inequality			
Gini	0.347	0.356	0.356	0.365	0.369	0.367	0.398
SD	0.717	0.726	0.723	0.753	0.752	0.748	0.781
IQR	0.944	0.991	1.000	1.046	1.051	1.035	1.085
90-50	0.715	0.744	0.744	0.767	0.798	0.792	0.904
50-10	1.051	1.062	1.024	1.076	1.039	1.046	1.012

C: Marginal Contribution to Change in Inequality

	Actual Change	Marginal Change Attributable to					Residual Change
		HH Types	NS	Attributes	Employment	Returns	
Gini	0.051	0.009	-0.0004	0.009	0.004	-0.002	0.031
	(14.8)	(17.6)	(-0.7)	(17.9)	(7.6)	(-3.5)	(61.1)
SD	0.064	0.009	-0.004	0.030	-0.001	-0.003	0.033
	(8.9)	(14.2)	(-6.2)	(47.3)	(-1.7)	(-5.3)	(51.6)
IQR	0.141	0.047	0.009	0.047	0.004	-0.016	0.050
	(14.9)	(33.1)	(6.6)	(33.2)	(3.0)	(-11.4)	(35.5)
90-50	0.188	0.029	0.0003	0.022	0.031	-0.006	0.112
	(26.3)	(15.4)	(0.2)	(11.8)	(16.5)	(-3.2)	(59.3)
50-10	-0.039	0.011	-0.037	0.052	-0.037	0.008	-0.034
	(-3.7)	(-27.5)	(96.7)	(-133.6)	(96.3)	(-20.1)	(88.3)
J-stat[a]	3.37	0.25	1.13	0.33	0.38	-0.35	1.64
	n/a	(7.4)	(33.4)	(9.7)	(11.2)	(-10.5)	(48.8)

Notes: In panel A, the percentiles are in constant (1999) dollar values. In panel C, the percentage change, relative to 1983-86, are in parentheses for the actual change; while, the percentage of the total change are in parentheses for the marginal change explained by each factor and the residual.

[a] The J-Stat in panel C is the Kullback-Leibler J statistic, which measures divergence between two distributions. See text for formula.

Table 5: Changes in Income Inequality 1983-86 to 1989-92 and 1989-92 to 1995-98

	Initial Period Distribution	Counterfactual Distribution Allowing Changes in					Final Period Distribution
		Household Structure	+National Superannuation	+Household Attributes	+Employment Outcomes	+Economic Returns	
		Changes between 1983-86 and 1989-92					
		A: Summary Measures of Income Inequality					
Gini	0.347	0.357	0.357	0.364	0.369	0.370	0.386
SD	0.717	0.724	0.726	0.746	0.750	0.750	0.763
IQR	0.944	0.987	0.993	1.025	1.033	1.043	1.062
90-50	0.715	0.750	0.752	0.768	0.802	0.811	0.865
50-10	1.051	1.060	1.066	1.108	1.067	1.069	1.066
		B: Marginal Contribution to Change in Inequality					
Gini	0.039	0.009	0.001	0.007	0.005	0.001	0.016
	(11.2)	(24.3)	(2.0)	(17.2)	(13.6)	(1.8)	(41.1)
SD	0.046	0.007	0.002	0.021	0.003	0.001	0.013
	(6.4)	(14.3)	(3.4)	(44.8)	(7.5)	(1.6)	(28.4)
IQR	0.118	0.043	0.006	0.032	0.008	0.010	0.020
	(12.5)	(36.3)	(4.8)	(27.1)	(6.9)	(8.4)	(16.6)
90-50	0.150	0.035	0.002	0.017	0.034	0.009	0.054
	(20.9)	(23.1)	(1.4)	(11.2)	(22.5)	(5.9)	(35.9)
50-10	0.015	0.009	0.006	0.042	-0.041	0.002	-0.003
	(1.4)	(59.1)	(42.2)	(287.5)	(-281.8)	(12.3)	(-19.4)
		Changes between 1989-92 and 1995-98					
		A: Summary Measures of Income Inequality					
Gini	0.386	0.385	0.384	0.390	0.388	0.384	0.398
SD	0.763	0.765	0.760	0.778	0.773	0.764	0.781
IQR	1.062	1.065	1.070	1.105	1.103	1.067	1.085
90-50	0.865	0.860	0.860	0.867	0.861	0.861	0.904
50-10	1.066	1.070	1.010	1.043	1.037	1.009	1.012
		B: Marginal Contribution to Change in Inequality					
Gini	0.012	-0.001	-0.001	0.006	-0.002	-0.004	0.015
	(3.2)	(-6.4)	(-9.7)	(48.7)	(-16.2)	(-33.2)	(116.8)
SD	0.018	0.002	-0.006	0.019	-0.005	-0.010	0.018
	(2.3)	(11.2)	(-33.1)	(106.9)	(-28.1)	(-55.8)	(99.0)
IQR	0.022	0.003	0.004	0.035	-0.001	-0.037	0.018
	(2.1)	(14.3)	(19.6)	(157.3)	(-6.3)	(-165.0)	(80.2)
90-50	0.038	-0.005	0.001	0.007	-0.006	-0.0003	0.043
	(4.4)	(-13.3)	(1.4)	(17.2)	(-15.4)	(-0.8)	(110.7)
50-10	-0.053	0.004	-0.060	0.033	-0.007	-0.027	0.003
	(-5.0)	(-8.3)	(112.7)	(-62.6)	(12.9)	(51.3)	(-5.9)

Notes: In panel B the "Initial Period" column contains the actual change in inequality between the initial and final periods (together with the percentage change, relative to initial period inequality, in parentheses); the "counterfactual" columns present the marginal change explained by each counterfactual (together with the percentage of the total change in parentheses); the "Final Period" column presents the unexplained change in inequality (together with the percentage of the total change in parentheses).

Table 6: Changes in Income Inequality 1983-86 to 1995-98, by Household Types

	Total Change in Inequality	Relative Marginal Effect of Changes in				
		National Superannuation	Household Attributes	Employment Outcomes	Economic Returns	Unexplained Change
			A: Single Adult Over 60			
Gini	0.012 (4.2)	-82.1	---	---	---	182.1
SD	0.007 (1.4)	-187.8	---	---	---	287.8
IQR	-0.055 (-11.0)	33.0	---	---	---	67.0
90-50	-0.067 (-7.8)	49.5	---	---	---	50.5
50-10	-0.016 (-9.9)	154.9	---	---	---	-54.9
			B: Multiple Adults Over 60			
Gini	0.026 (8.7)	31.8	---	---	---	68.2
SD	0.053 (10.5)	26.5	---	---	---	73.4
IQR	-0.067 (-9.2)	-40.6	---	---	---	140.6
90-50	0.064 (7.6)	19.9	---	---	---	80.1
50-10	-0.011 (-3.1)	-201.8	---	---	---	301.8
			C: Single Adult Under 60			
Gini	0.062 (18.9)	-0.0	2.7	43.1	23.2	31.0
SD	0.054 (7.6)	-0.1	50.5	113.4	24.4	-85.2
IQR	0.276 (37.0)	0.0	21.6	45.4	35.7	-2.6
90-50	0.182 (28.2)	0.0	-23.5	25.6	39.4	58.5
50-10	-0.030 (-2.8)	0.0	-309.9	91.3	-27.7	346.3
			D: Multiple Adults Under 60			
Gini	0.054 (19.8)	-0.3	3.8	30.1	12.6	53.8
SD	0.124 (21.0)	-0.3	24.8	25.9	10.8	38.7
IQR	0.134 (21.6)	-2.3	-1.1	34.8	7.3	61.3
90-50	0.125 (22.5)	-0.2	-16.6	29.5	8.8	7.9
50-10	0.223 (32.8)	-0.1	24.2	28.8	30.8	16.3
			E: Single Adult With Children			
Gini	-0.041 (-13.9)	0.2	-11.7	66.7	1.6	43.3
SD	-0.047 (-8.5)	-0.1	-51.7	78.8	-15.0	88.0
IQR	-0.085 (-18.7)	-0.9	-61.4	110.4	-60.2	112.0
90-50	-0.250 (-32.6)	0.0	-9.4	66.0	39.2	4.2
50-10	-0.020 (-4.8)	0.0	-278.6	166.2	-170.1	382.5
			F: Multiple Adults With Children			
Gini	0.060 (21.0)	0.0	12.3	15.8	-4.5	76.4
SD	0.081 (12.3)	-0.1	30.2	-15.6	-2.0	87.5
IQR	0.150 (23.2)	-0.6	12.2	34.3	-27.0	81.1
90-50	0.149 (25.3)	-0.0	18.9	24.2	10.7	46.3
50-10	0.074 (10.8)	0.1	30.1	8.0	-36.7	98.5

Notes: In the "Total Change" column, numbers are raw changes (with percentage changes in parentheses); the relative marginal effects in other columns are percentages of the total change.

Table A1(a): Counterfactual Analysis for Single Adult Under 60 Households

Variable	Attributes	Employment Outcomes		Economic Returns	
		1983-86	1995-98	1983-86	1995-98
	(1)	(2)	(3)	(4)	(5)
Female	-0.041	-1.139	-0.824	-0.322	-0.212
	(.110)	(.011)	(.009)	(.045)	(.067)
Maori	0.927	-0.371	-0.848	-0.226	-0.243
	(.234)	(.019)	(.014)	(.112)	(.123)
Aged 15-18	1.121	-1.985	0.949	-0.175	0.092
	(.724)	(.066)	(.060)	(.226)	(.438)
Aged 25-39	0.905	-0.871	0.336	0.260	0.502
	(.210)	(.042)	(.020)	(.074)	(.146)
Aged 40-54	1.396	-1.614	0.197	0.281	0.637
	(.211)	(.041)	(.020)	(.076)	(.146)
Aged 55-59	1.192	-2.502	-0.629	0.080	0.389
	(.230)	(.042)	(.021)	(.083)	(.160)
Highest Qualification					
School	0.695	0.240	1.064	0.348	0.448
	(.146)	(.014)	(.011)	(.058)	(.089)
Vocational	0.385	0.609	0.914	0.426	0.441
	(.142)	(.014)	(.011)	(.056)	(.090)
University	0.917	1.665	1.660	0.686	0.880
	(.170)	(.023)	(.014)	(.071)	(.101)
Intercept	-1.214	3.088	0.563	9.901	9.487
	(.223)	(.042)	(.020)	(.079)	(.154)
Psuedo R^2	0.043	0.165	0.124	0.151	0.139
Number of Observations	1,608	814	794	794	781

Notes: Standard errors are in parentheses. All models are estimated using sampling weights. In column (1), the dependent variable equals 1 if the observation is from 1995-98, and 0 if from 1983-86, and the specification estimated is a Logit model. In columns (2) and (3), the dependent variable is defined in table 2, and the specifications are Logit models. In columns (4) and (5), the dependent variable is log(household income), and the specifications are estimated using Median regressions. See text for further details.

Table A1(b): Counterfactual Analysis for Multiple Adults Under 60 Households

Variable	Attributes	Employment Outcomes		Economic Returns	
		1983-86	1995-98	1983-86	1995-98
	(1)	(2)	(3)	(4)	(5)
Number of Adults	0.193	0.954	0.722	0.246	0.292
	(.046)	(.006)	(.004)	(.021)	(.020)
Fraction of Adults:					
Female	0.143	-0.415	-0.630	-0.341	-0.108
	(.173)	(.016)	(.011)	(.074)	(.073)
Maori	1.506	0.010	-0.504	-0.131	-0.121
	(.180)	(.011)	(.010)	(.099)	(.066)
Married	0.210	0.266	0.625	0.127	0.146
	(.089)	(.008)	(.006)	(.034)	(.041)
Aged 18-25	-0.338	0.329	-0.245	0.095	-0.215
	(.287)	(.029)	(.020)	(.120)	(.124)
Aged 25-39	0.886	0.792	1.052	0.526	0.442
	(.124)	(.012)	(.010)	(.053)	(.058)
Aged 40-54	1.792	-0.590	0.242	0.525	0.434
	(.133)	(.012)	(.010)	(.057)	(.061)
Aged 55-59	1.304	-1.740	-0.818	0.482	0.233
	(.160)	(.014)	(.011)	(.067)	(.073)
Aged 60+	1.279	-3.924	-2.236	0.333	0.098
	(.258)	(.019)	(.017)	(.099)	(.121)
Fraction of Adults with Highest Qualifications					
School	1.554	0.512	0.943	0.338	0.364
	(.118)	(.009)	(.008)	(.052)	(.054)
Vocational	1.158	0.826	0.958	0.353	0.440
	(.117)	(.010)	(.008)	(.047)	(.054)
University	1.892	0.157	0.424	0.479	0.591
	(.140)	(.011)	(.009)	(.063)	(.059)
Intercept	-2.537	---	---	9.866	9.651
	(.217)			(.092)	(.099)
First Break	---	-1.359	0.119	---	---
		(.022)	(.016)		
Second Break	---	1.378	2.162	---	---
		(.022)	(.016)		
Psuedo R^2	0.081	0.135	0.091	0.130	0.135
Number of Observations	5,185	2,802	2,383	2,772	2,341

Notes: Standard errors are in parentheses. All models are estimated using sampling weights. In column (1), the dependent variable equals 1 if the observation is from 1995-98, and 0 if from 1983-86, and the specification estimated is a Logit model. In columns (2) and (3), the dependent variable is defined in table 2, and the specifications are Ordered Logit models. In columns (4) and (5), the dependent variable is log(household income), and the specifications are estimated using Median regressions. See text for further details.

Table A1(c): Counterfactual Analysis for Single Adult With Children Households

Variable	Attributes	Employment Outcomes		Economic Returns	
		1983-86	1995-98	1983-86	1995-98
	(1)	(2)	(3)	(4)	(5)
Female	-0.191	-2.339	-0.700	0.045	-0.098
	(.230)	(.023)	(.015)	(.037)	(.055)
Maori	1.028	-0.655	-0.512	-0.010	-0.062
	(.185)	(.025)	(.016)	(.030)	(.041)
Married	1.155	-0.317	-1.582	0.766	0.080
	(.436)	(.048)	(.043)	(.049)	(.078)
Aged 25-39	1.266	0.712	-0.544	-0.071	-0.019
	(.628)	(.130)	(.067)	(.085)	(.152)
Aged 40-54[(a)]	2.271	-0.094	-1.472	0.016	0.030
	(.747)	(.137)	(.072)	(.101)	(.181)
Aged 55-59	2.692	---	---	-0.497	-0.707
	(1.48)			(.160)	(.329)
Aged 60+	-1.953	---	---	-0.314	-0.113
	(1.62)			(.155)	(.388)
Highest Qualification					
School	0.614	0.630	1.272	0.012	0.023
	(.166)	(.020)	(.017)	(.023)	(.042)
Vocational	0.849	1.207	1.239	0.065	0.086
	(.196)	(.022)	(.017)	(.029)	(.048)
University	0.914	2.043	1.805	0.410	0.184
	(.296)	(.031)	(.022)	(.036)	(.068)
Fraction of Persons:					
Aged 0-4	3.385	-5.590	-7.117	0.139	0.810
	(.992)	(.1374	(.105)	(.133)	(.244)
Aged 5-14	2.130	-0.547	-3.349	0.429	0.975
	(.793)	(.145)	(.081)	(.105)	(.199)
Aged 15-18	1.718	1.822	-1.560	0.612	1.332
	(.928)	(.153)	(.090)	(.123)	(.240)
Intercept	-1.794	1.259	0.943	9.591	9.487
	(.715)	(.140)	(.072)	(.098)	(.179)
Psuedo R^2	0.060	0.319	0.200	0.135	0.110
Number of Observations	1,013	428	585	423	576

Notes: Standard errors are in parentheses. All models are estimated using sampling weights. In column (1), the dependent variable equals 1 if the observation is from 1995-98, and 0 if from 1983-86, and the specification estimated is a Logit model. In columns (2) and (3), the dependent variable is defined in table 2, and the specifications are Logit models. In columns (4) and (5), the dependent variable is log(household income), and the specifications are estimated using Median regressions. See text for further details.

[(a)] Aged 40+ in columns (2) and (3).

Table A1(d): Counterfactual Analysis for Multiple Adults With Children Households

Variable	Attributes	Employment Outcomes		Economic Returns	
		1983-86	1995-98	1983-86	1995-98
	(1)	(2)	(3)	(4)	(5)
Number of Adults	-0.027	2.093	1.311	0.389	0.271
	(.056)	(.007)	(.005)	(.023)	(.029)
Fraction of Adults:					
Female	0.939	-2.322	-1.589	0.080	-0.204
	(.255)	(.024)	(.019)	(.083)	(.126)
Maori	0.843	-0.137	-0.083	-0.026	-0.023
	(.103)	(.008)	(.008)	(.036)	(.052)
Married	-0.207	0.419	1.139	0.173	0.020
	(.128)	(.010)	(.010)	(.035)	(.063)
Aged 15-18	0.249	-2.076	1.566	---	---
	(.425)	(.043)	(.032)		
Aged 25-39	0.986	0.718	1.242	0.665	0.371
	(.280)	(.025)	(.020)	(.087)	(.135)
Aged 40-54	2.196	0.413	1.423	0.671	0.615
	(.314)	(.027)	(.023)	(.098)	(.156)
Aged 55-59	0.621	-2.912	-1.902	0.120	0.043
	(.558)	(.052)	(.043)	(.206)	(.297)
Aged 60+	3.168	-7.877	-5.965	-0.465	-0.192
	(.554)	(.038)	(.041)	(.123)	(.252)
Fraction of Adults with Highest Qualifications					
School	1.744	0.598	0.759	0.235	0.327
	(.099)	(.008)	(.007)	(.025)	(.052)
Vocational	1.109	0.409	0.809	0.325	0.442
	(.103)	(.008)	(.007)	(.026)	(.054)
University	2.035	0.414	0.374	0.611	0.725
	(.117)	(.009)	(.008)	(.032)	(.058)
Fraction of Persons:					
Aged 0-4	1.006	-5.015	-2.830	0.090	0.005
	(.401)	(.035)	(.029)	(.119)	(.199)
Aged 5-14	0.676	-0.611	-0.811	0.466	0.432
	(.325)	(.029)	(.024)	(.101)	(.163)
Aged 15-18	0.013	3.102	-2.437	0.639	0.261
	(.534)	(.049)	(.040)	(.126)	(.234)
Intercept	-2.382	---	---	9.047	9.669
	(.379)			(.153)	(.190)
First Break	---	-0.634	1.191	---	---
		(.035)	(.029)		
Second Break	---	3.913	4.372	---	---
		(.035)	(.029)		
Psuedo R^2	0.079	0.222	0.126	0.176	0.129
Number of Observations	6,741	4,008	2,733	3,931	2,672

Notes: Standard errors are in parentheses. All models are estimated using sampling weights. In column (1), the dependent variable equals 1 if the observation is from 1995-98, and 0 if from 1983-86, and the specification estimated is a Logit model. In columns (2) and (3), the dependent variable is defined in table 2, and the specifications are Ordered Logit models. In columns (4) and (5), the dependent variable is log(household income),

Figure 1: Relative Changes in Household Income Levels and Inequality, 1983/84 -- 1997/98

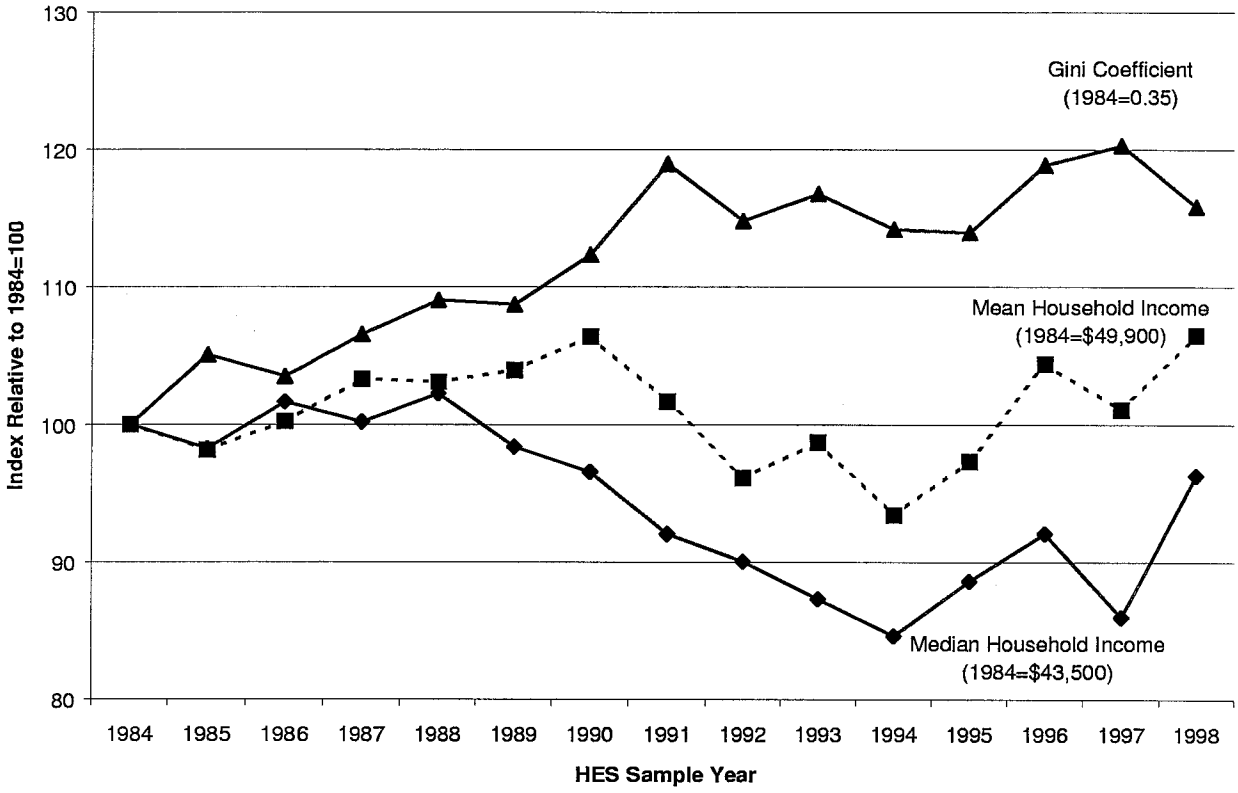

Figure 2a: Distribution of Households by Type, 1983/84 -- 1997/98

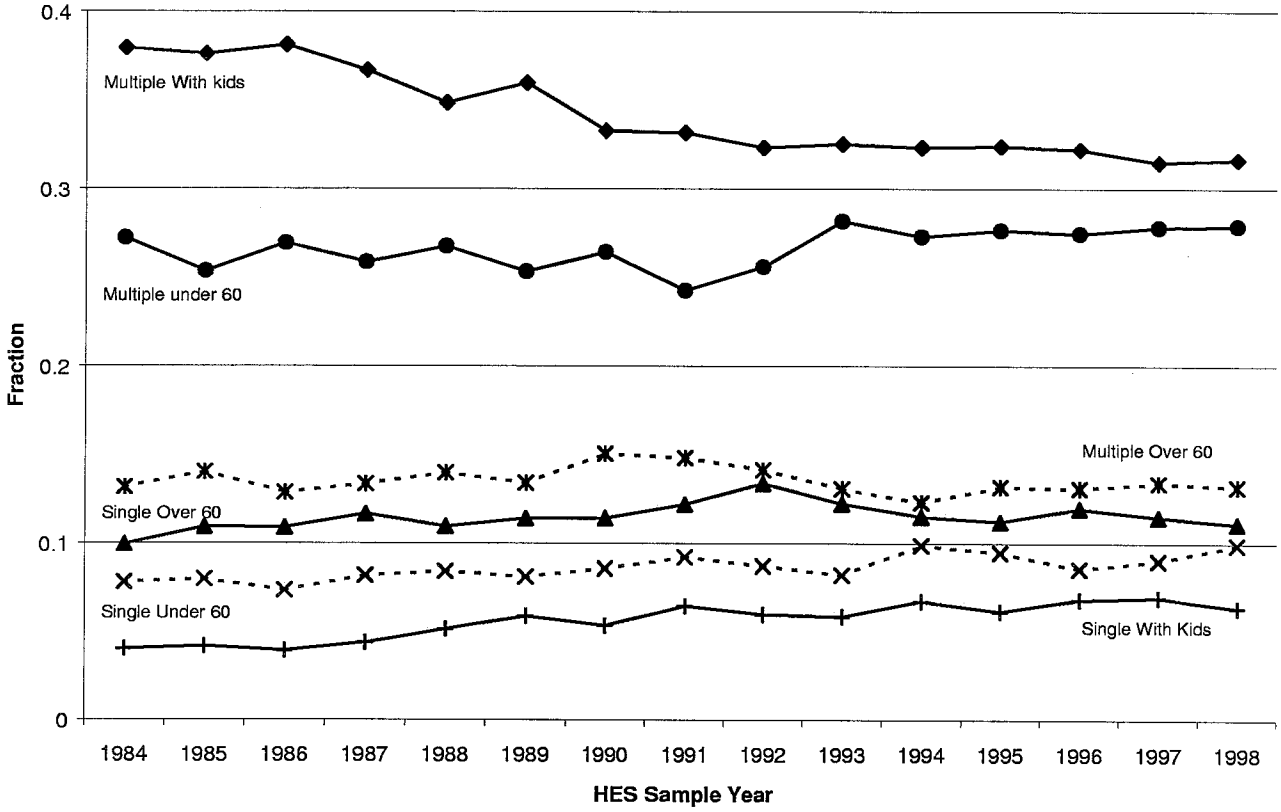

Figure 2b: Fraction of Adults Employed Fulltime, 1983/84 -- 1997/98

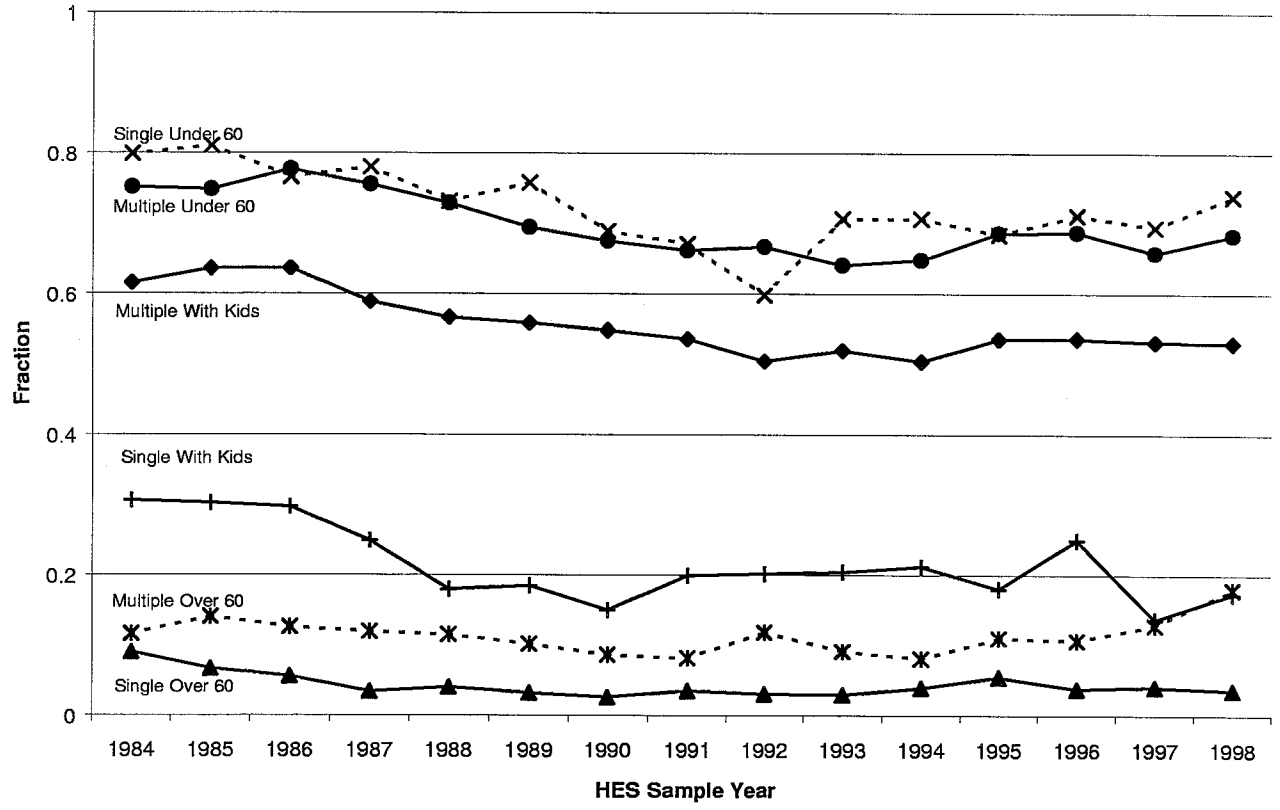

Figure 3a: Distribution of Household Gross Incomes: 1983-86 and 1995-98

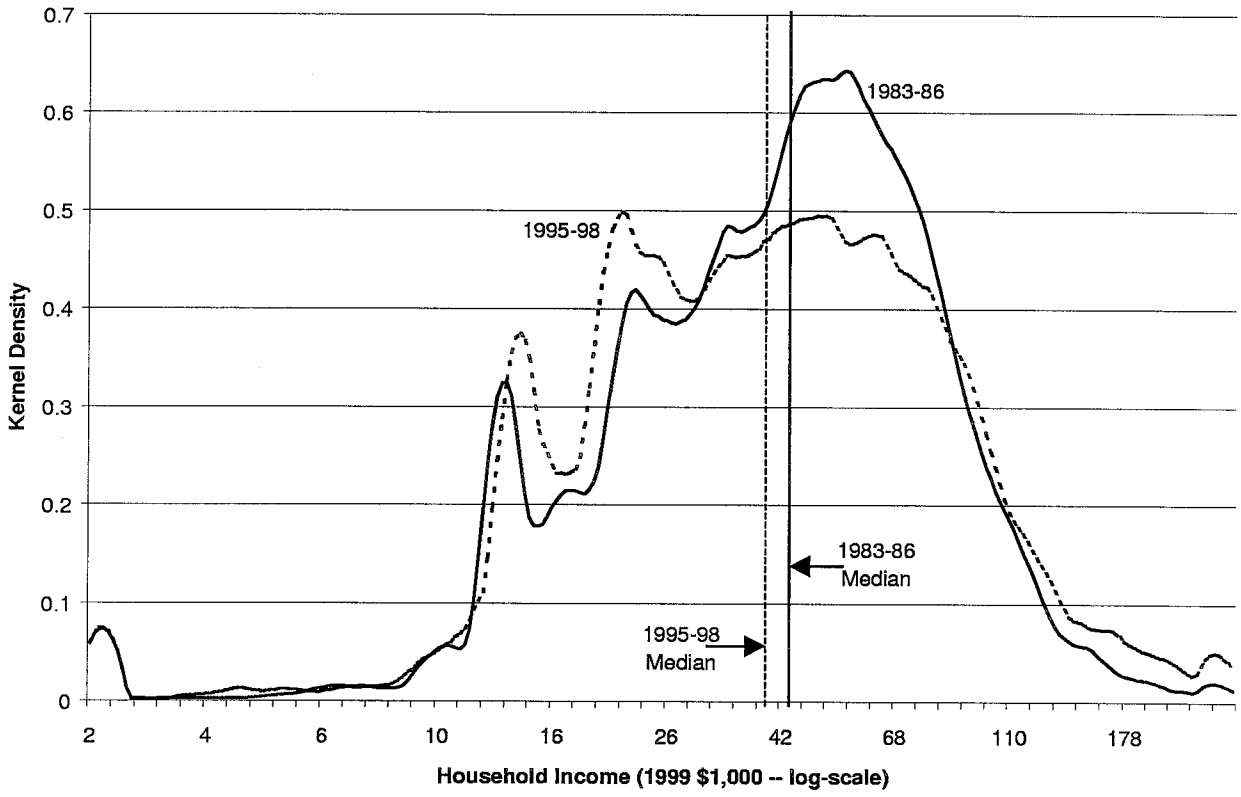

Figure 3b: Changes in Distribution Between 1983-86 and 1995-98

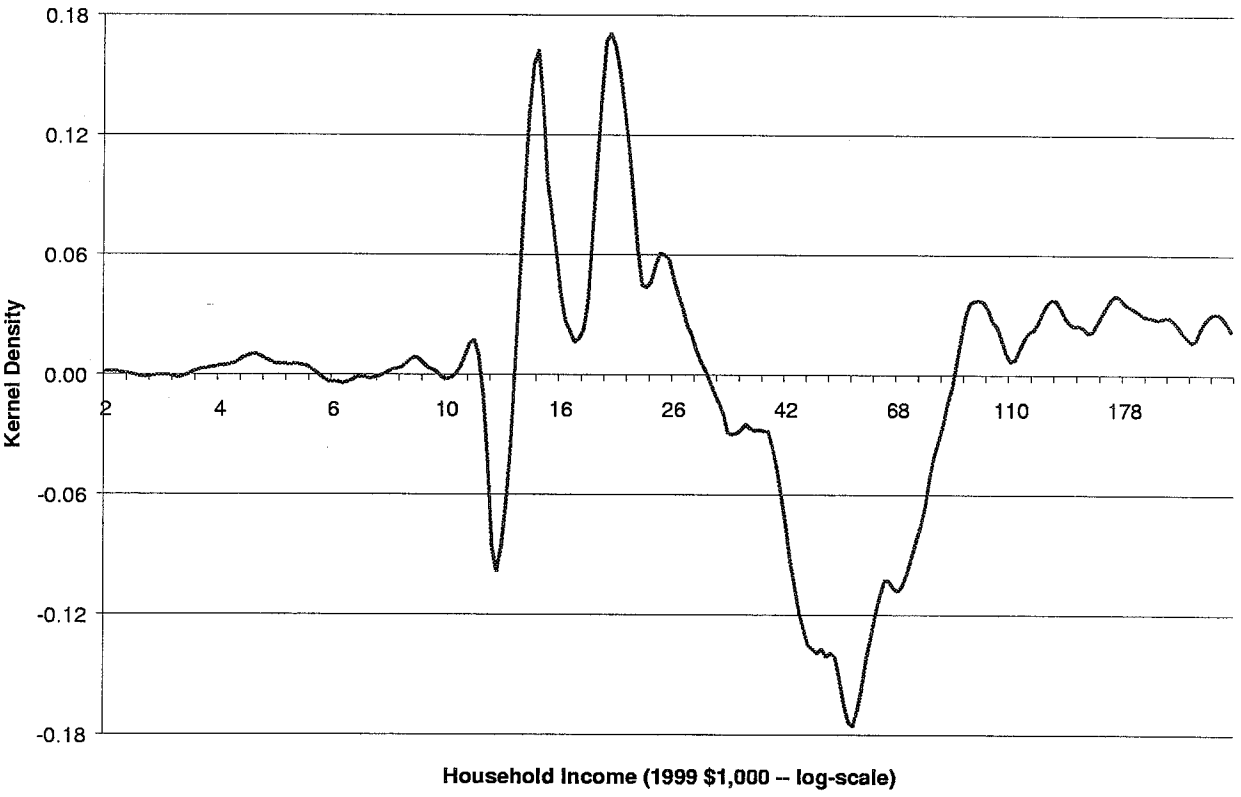

Figure 4a: Distribution of Household Incomes 1983-86: Contributions by Household Type

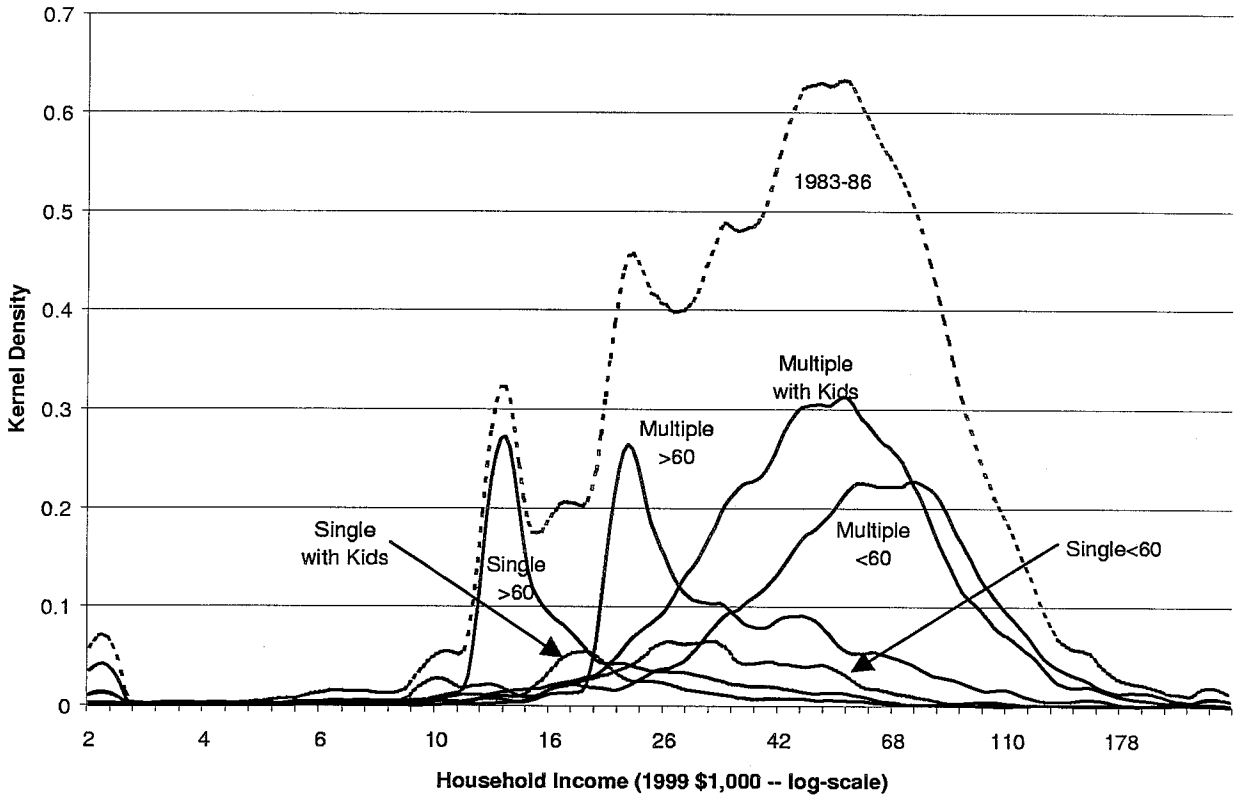

Figure 4b: Distribution of Household Incomes 1995-98: Contributions by Household Type

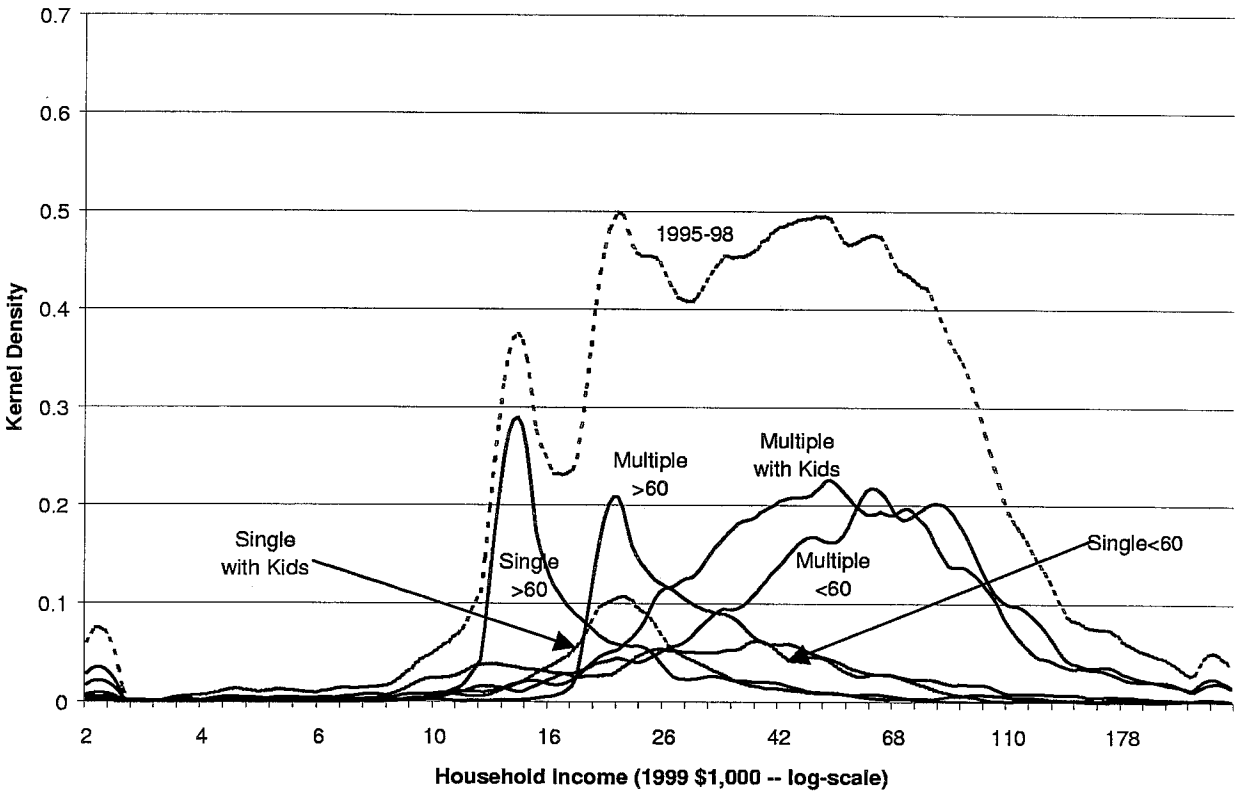

Figure 5a: Distribution of Household Incomes: Household Type Changes

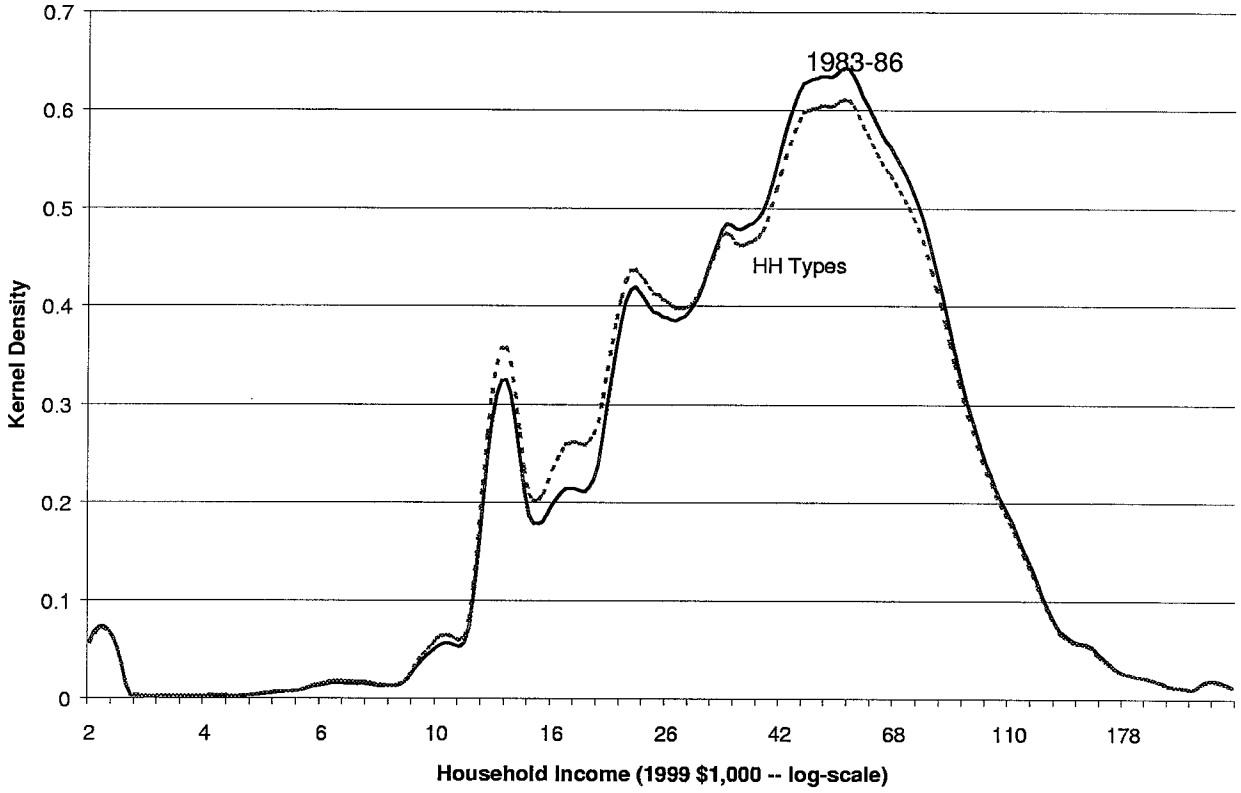

Figure 5b: Changes Between 1983-86 and 1995-98: Household Type Contribution

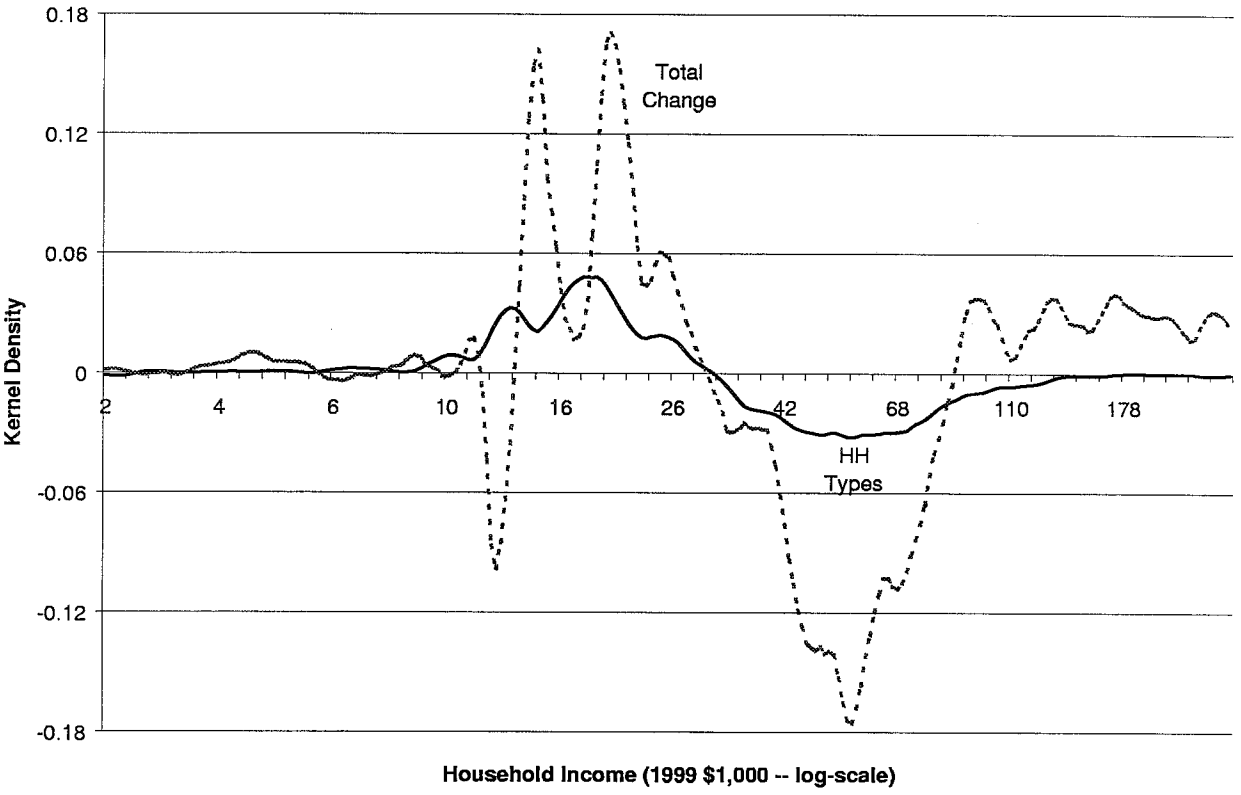

Figure 6a: Distribution of Household Incomes: National Superannuation Rate Changes

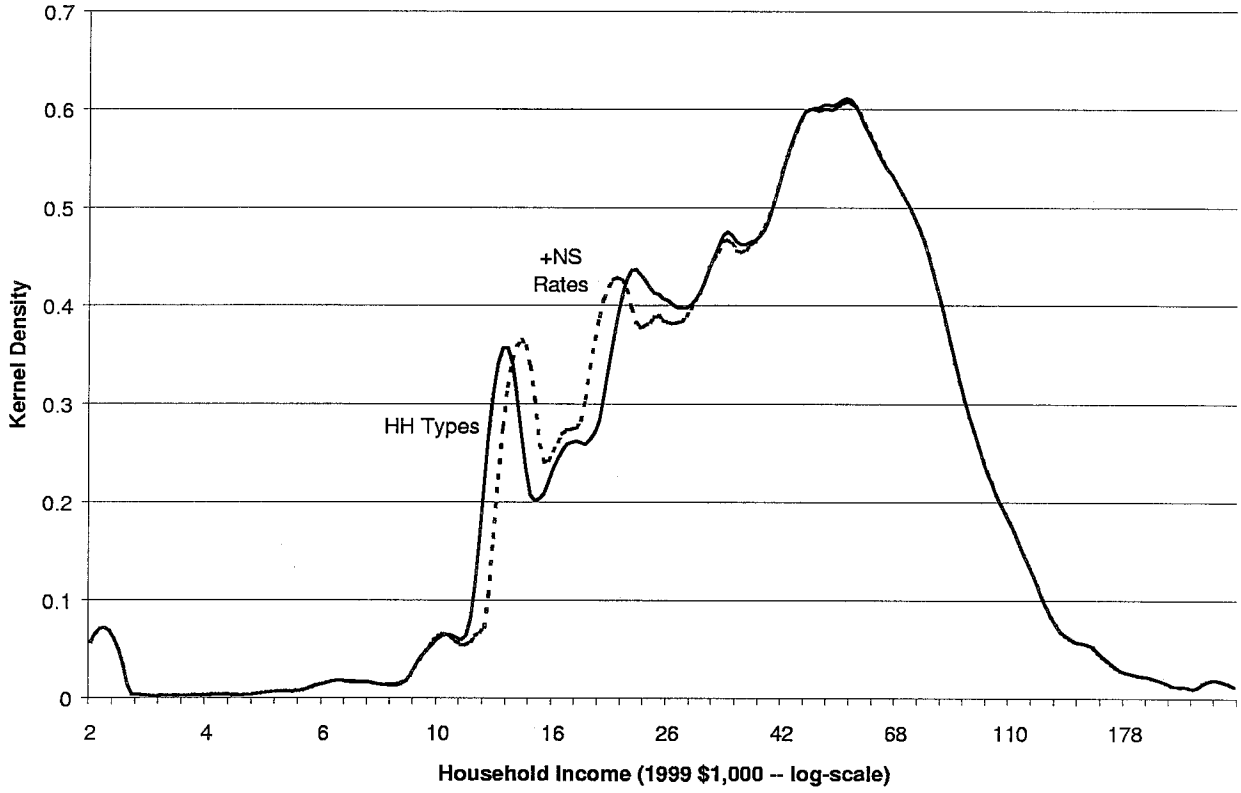

Figure 6b: Changes Between 1983-86 and 1995-98: National Super Rate Contribution

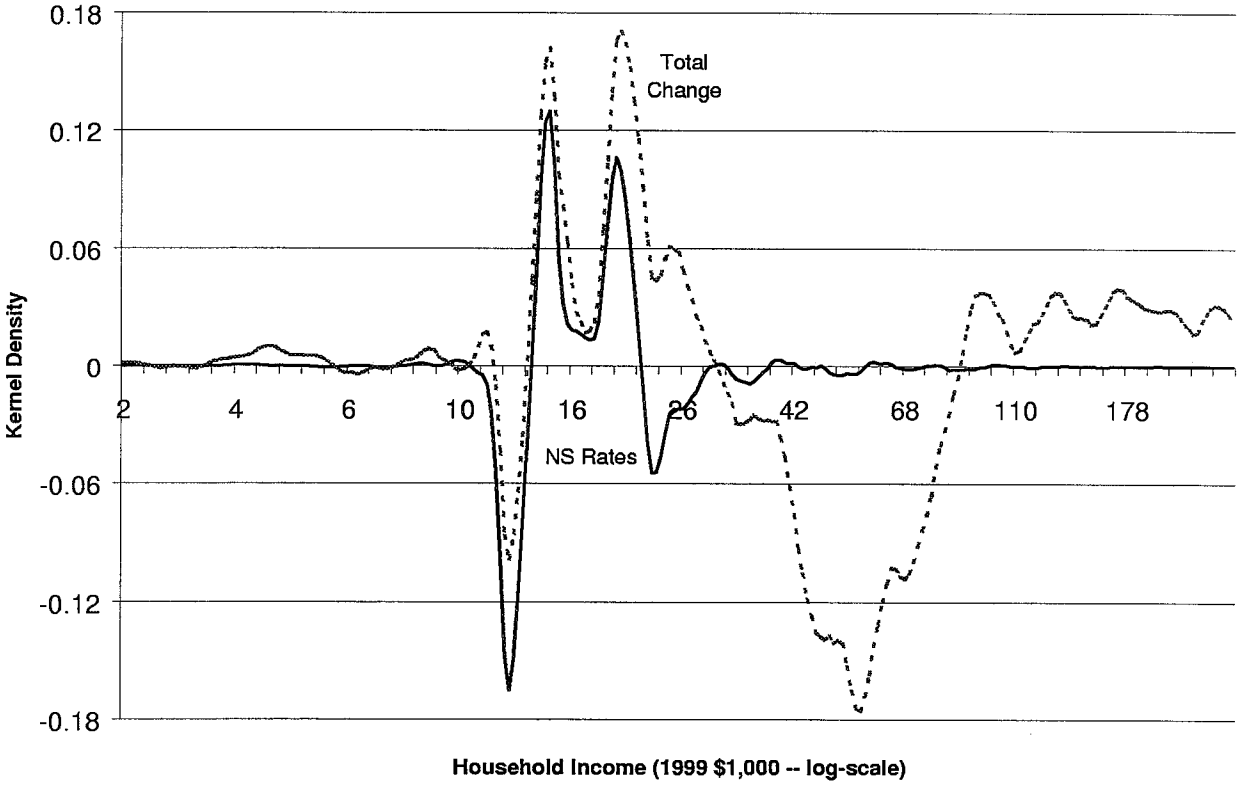

Figure 7a: Distribution of Household Incomes: Sociodemographic Attribute Changes

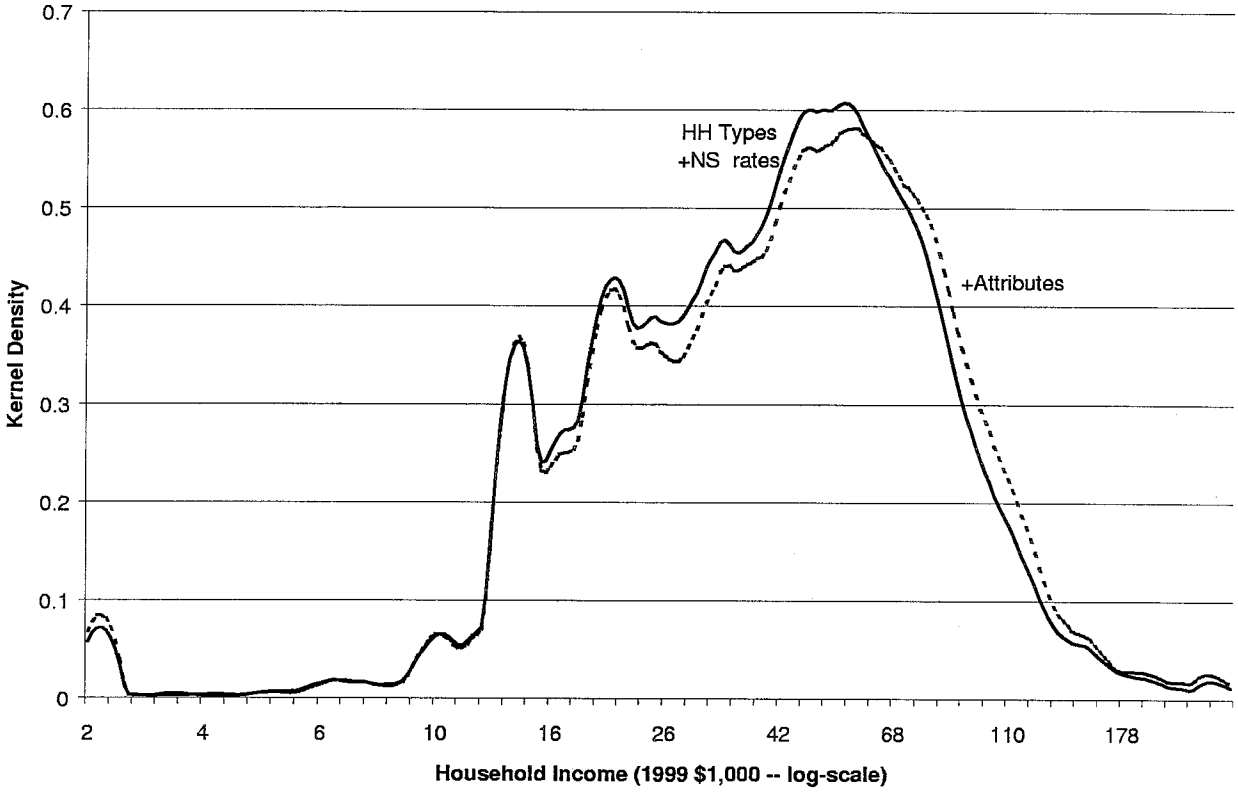

Figure 7b: Changes Between 1983-86 and 1995-98: Attributes Contribution

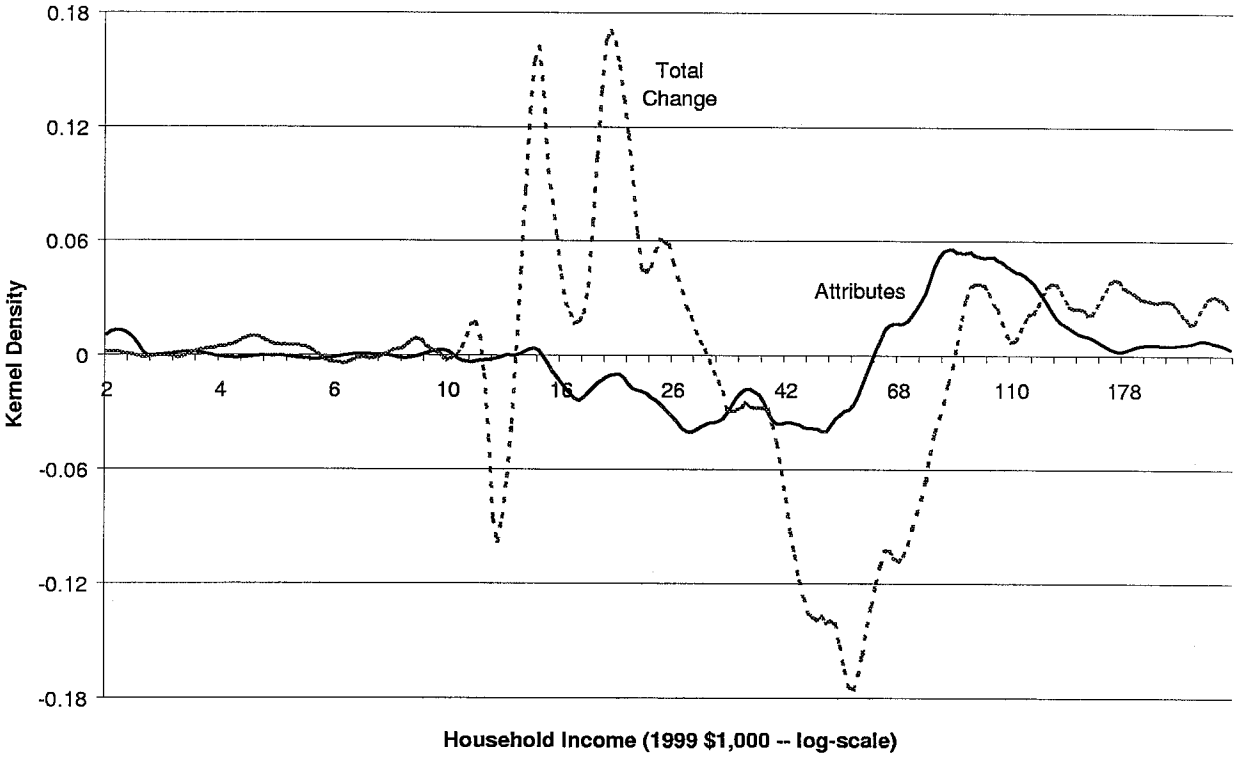

Figure 8a: Distribution of Household Incomes: Employment Changes

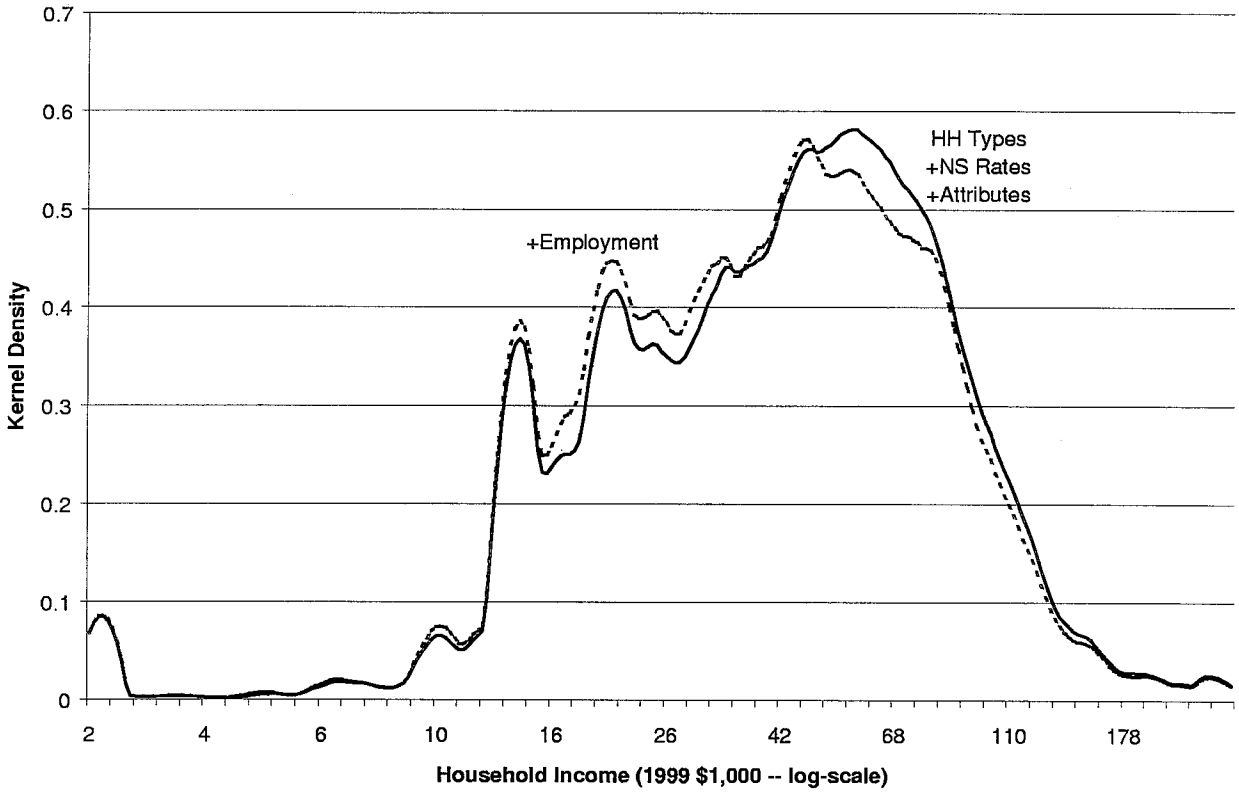

Figure 8b: Changes Between 1983-86 and 1995-98: Employment Contribution

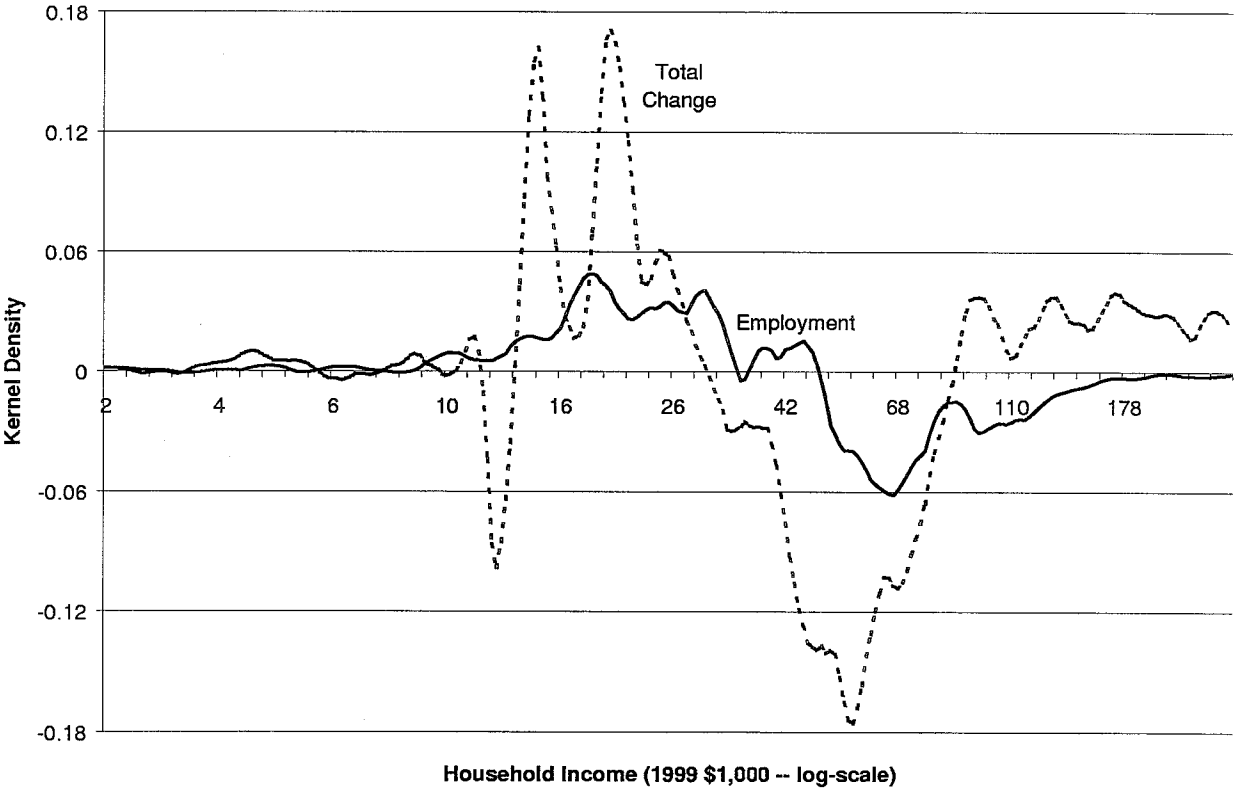

Figure 9a: Distribution of Household Incomes: Changes in Returns to Attributes

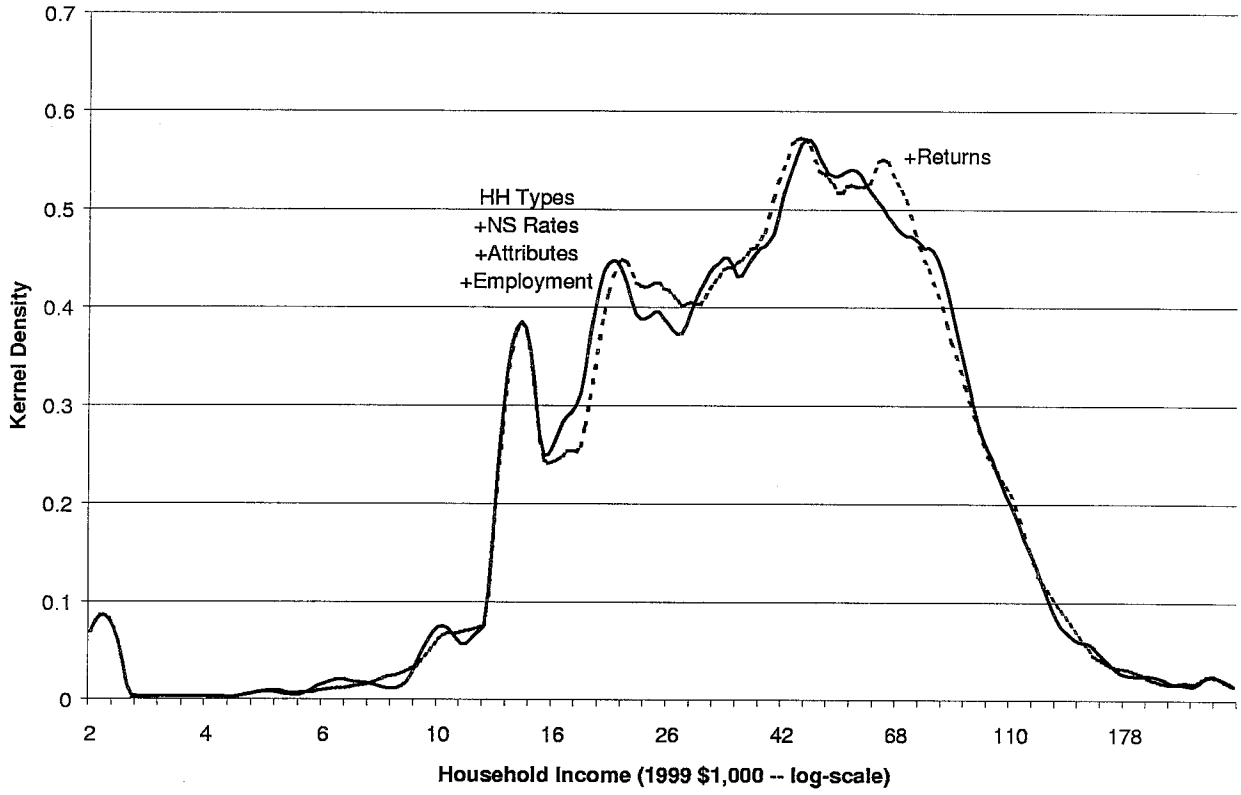

Figure 9b: Changes Between 1983-86 and 1995-98: Returns Contribution

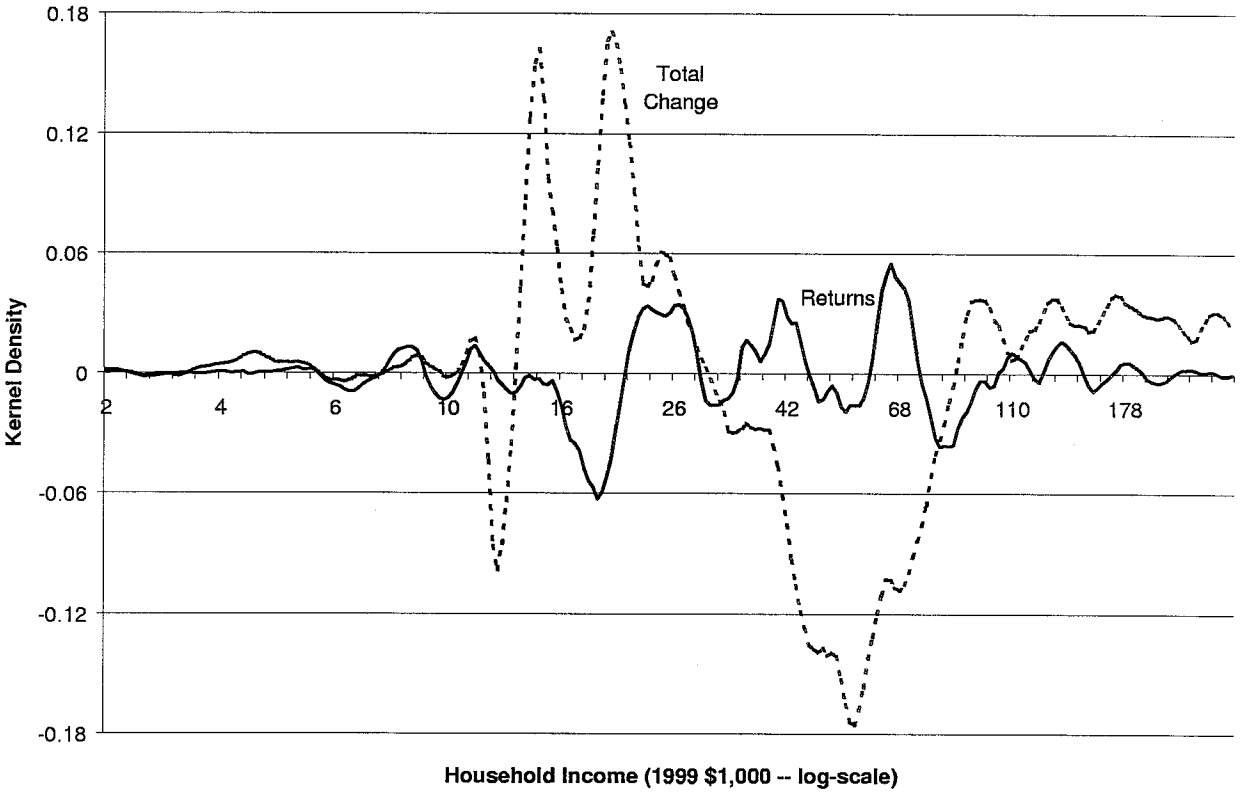

Figure 10a: Distribution of Household Incomes: Explained Changes

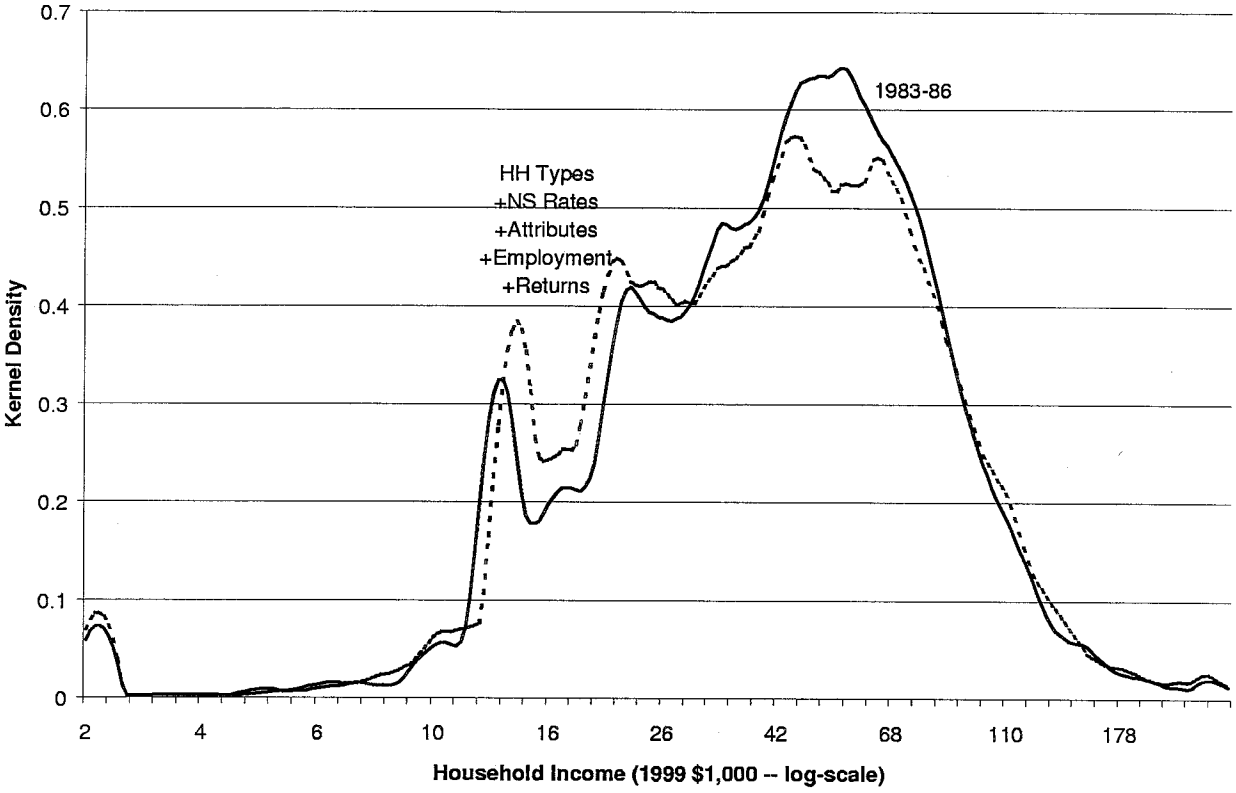

Figure 10b: Changes Between 1983-86 and 1995-98: Explained Contribution

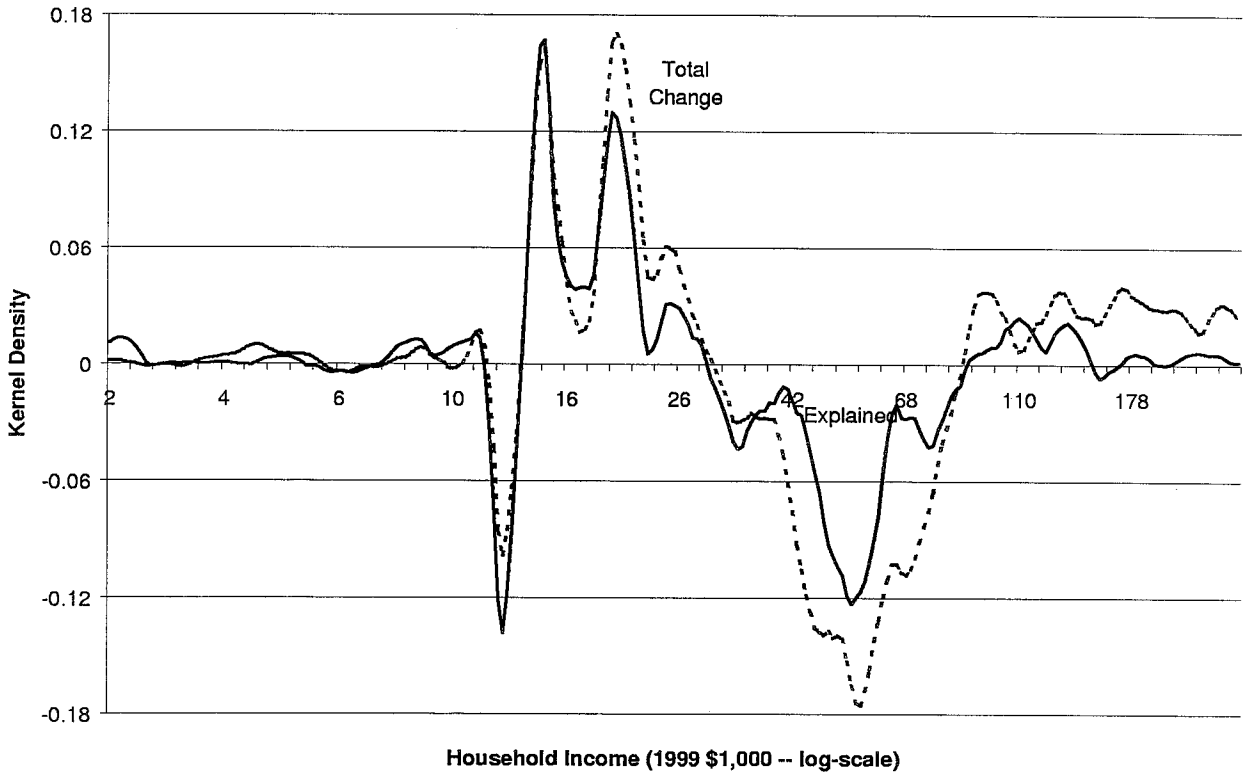

Figure 11a: Distribution of Household Incomes: Unexplained Changes

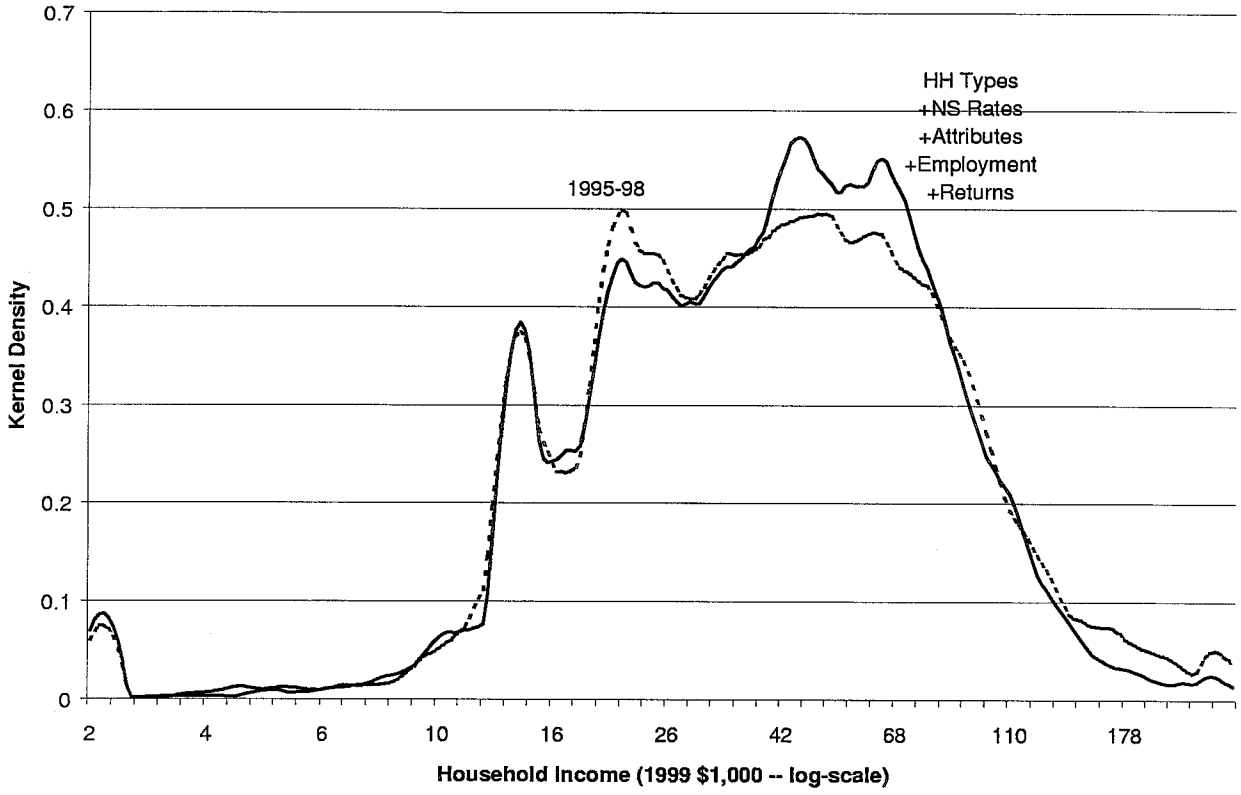

Figure 11b: Changes Between 1983-86 and 1995-98: Unexplained Contribution

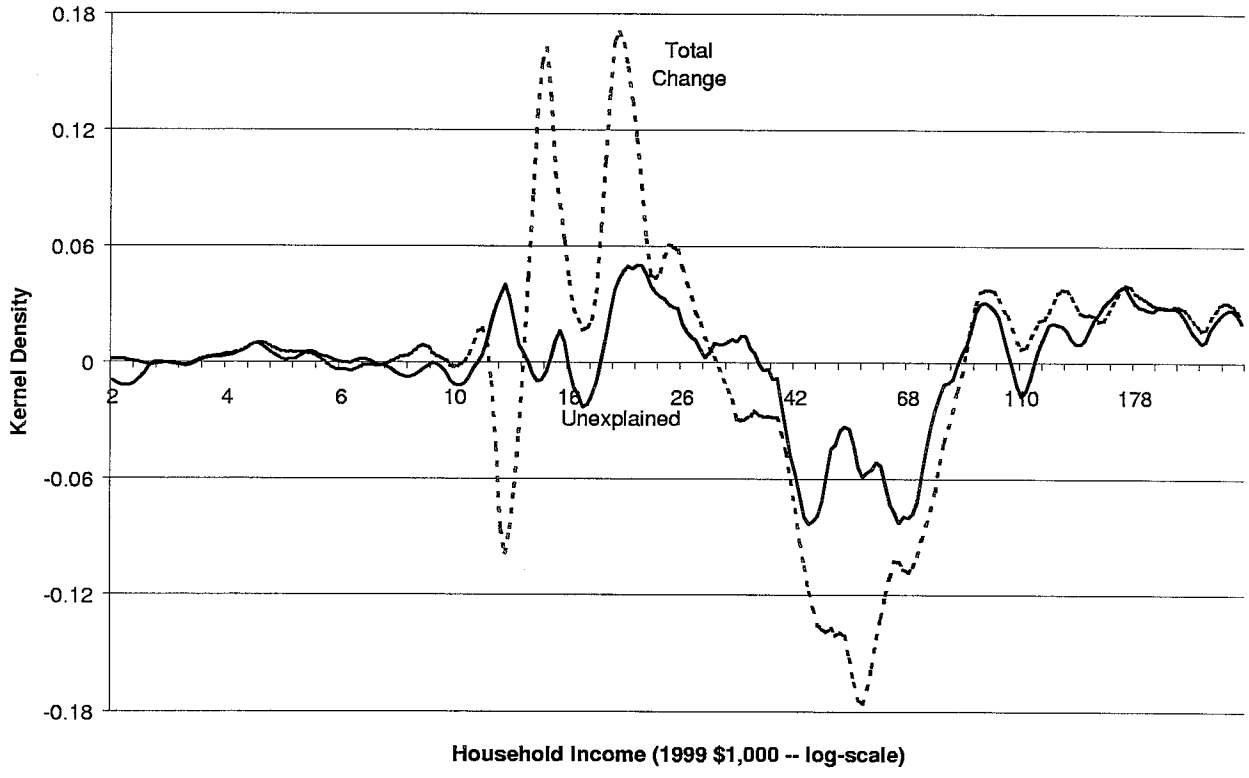

Figure 12a: Distribution of Household Gross Incomes: 1983-86 and 1989-92

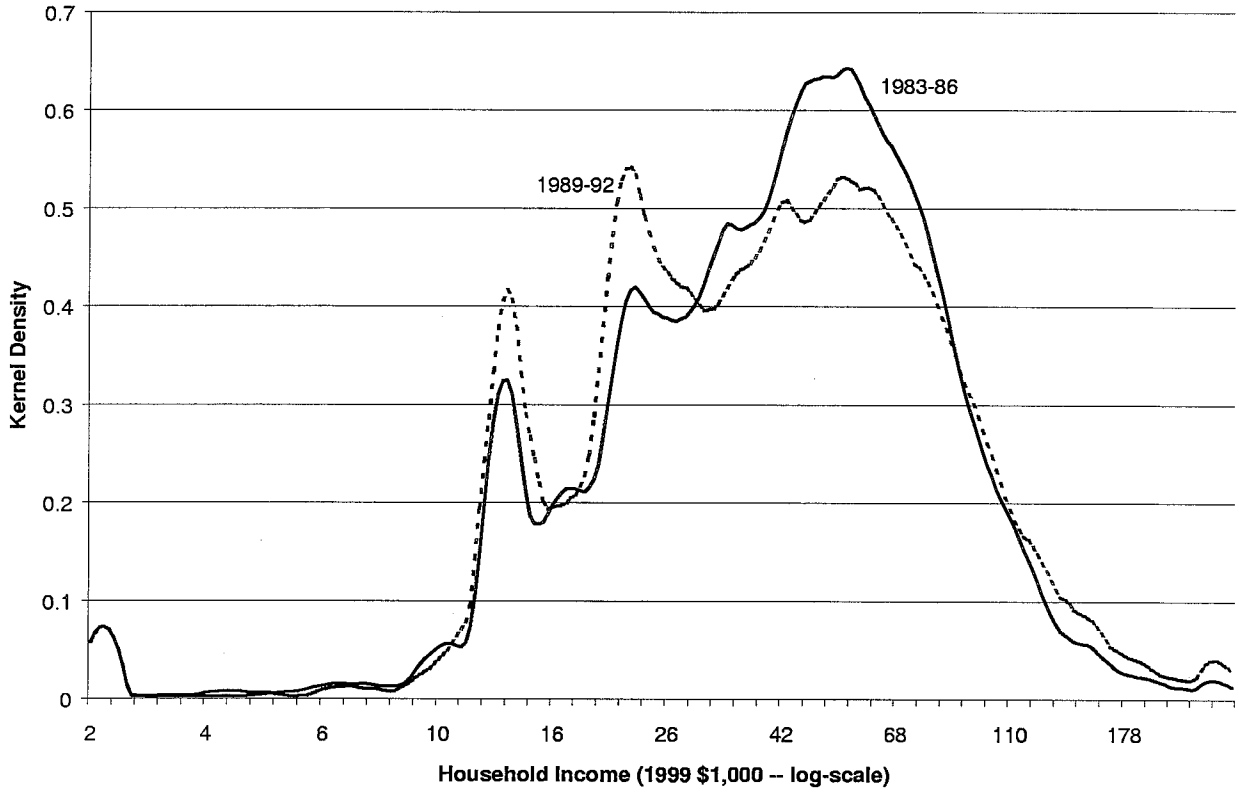

Figure 12b: Distribution of Household Gross Incomes: 1989-92 and 1995-98

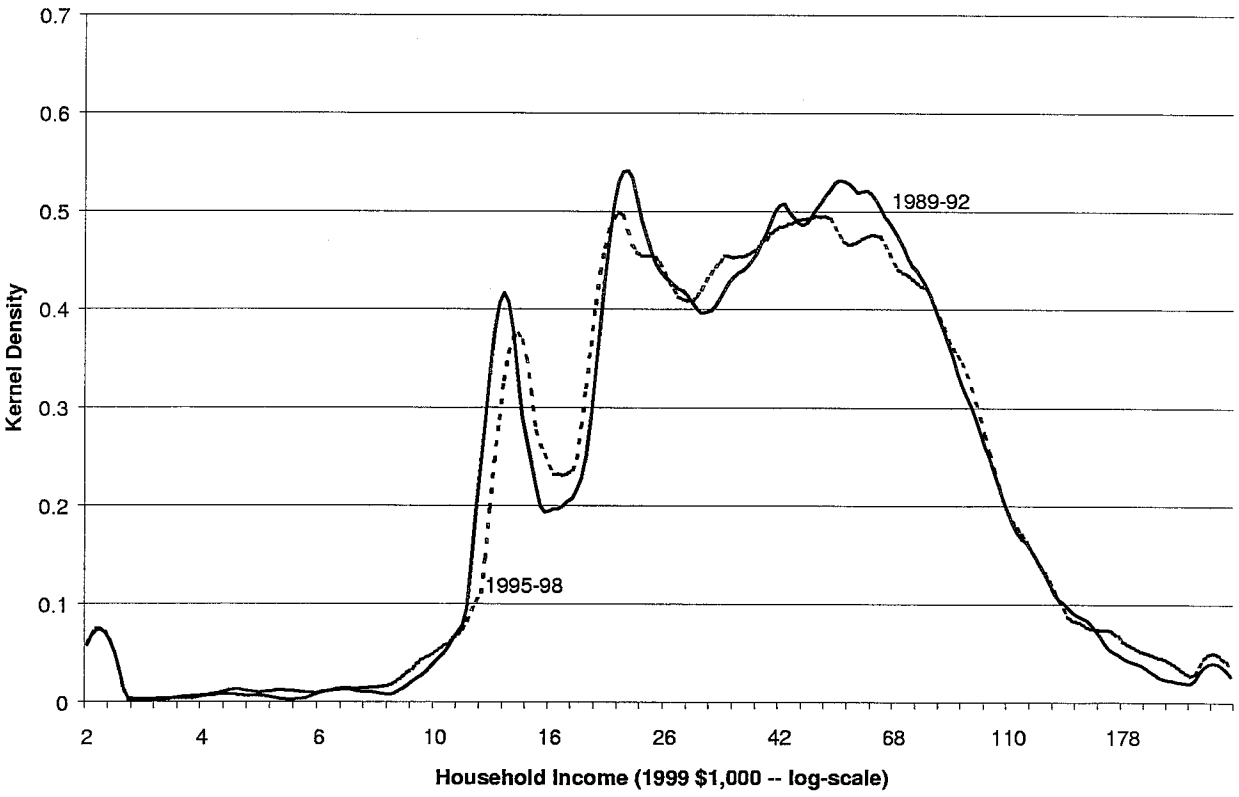

Figure 13a: Distribution of Household Incomes Changes: 1983-86 to 1989-92

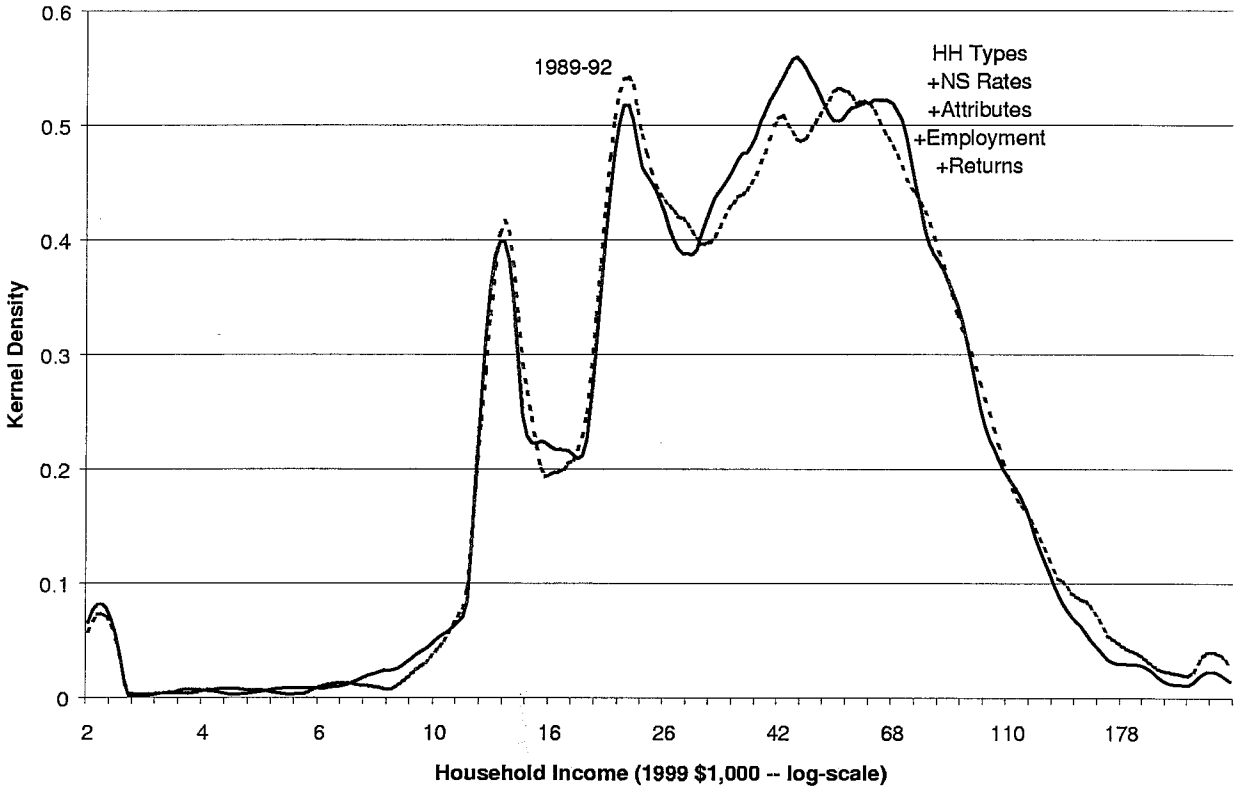

Figure 13b: Changes Between 1983-86 and 1989-92: Explained Contribution

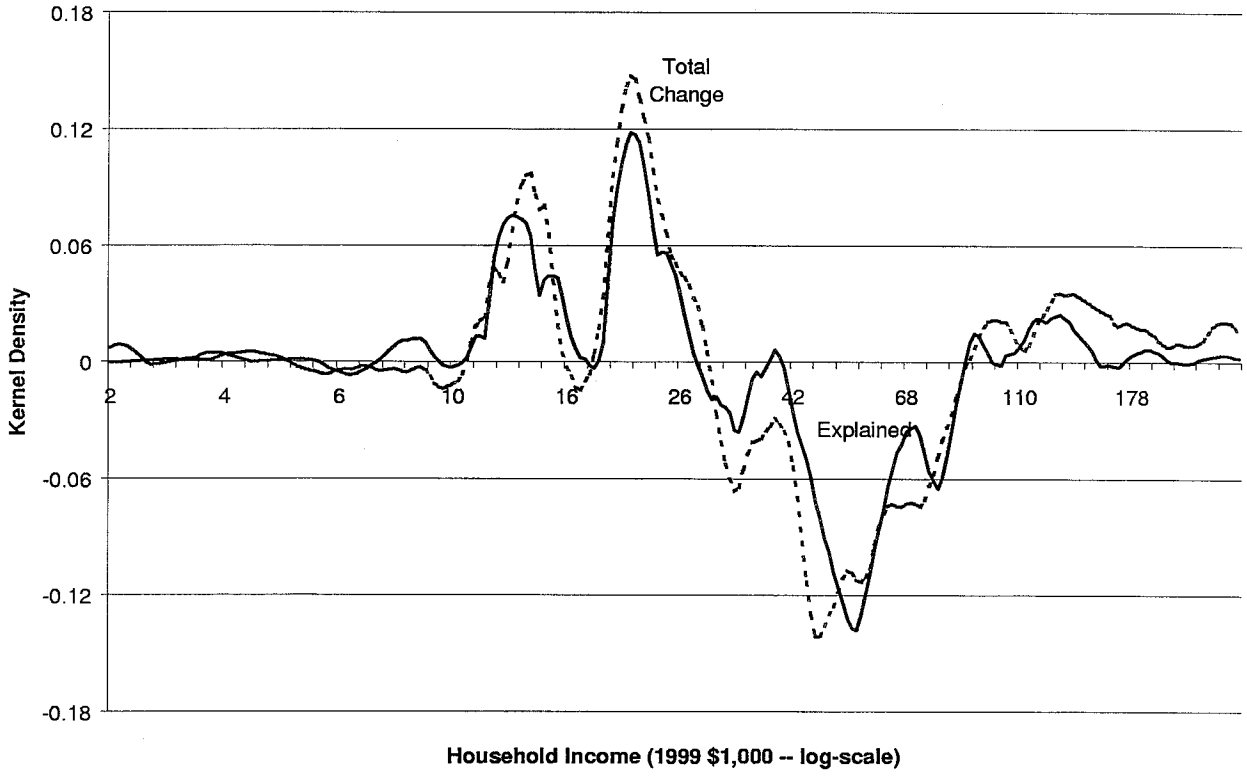

Figure 14a: Distribution of Household Incomes Changes: 1989-92 to 1995-98

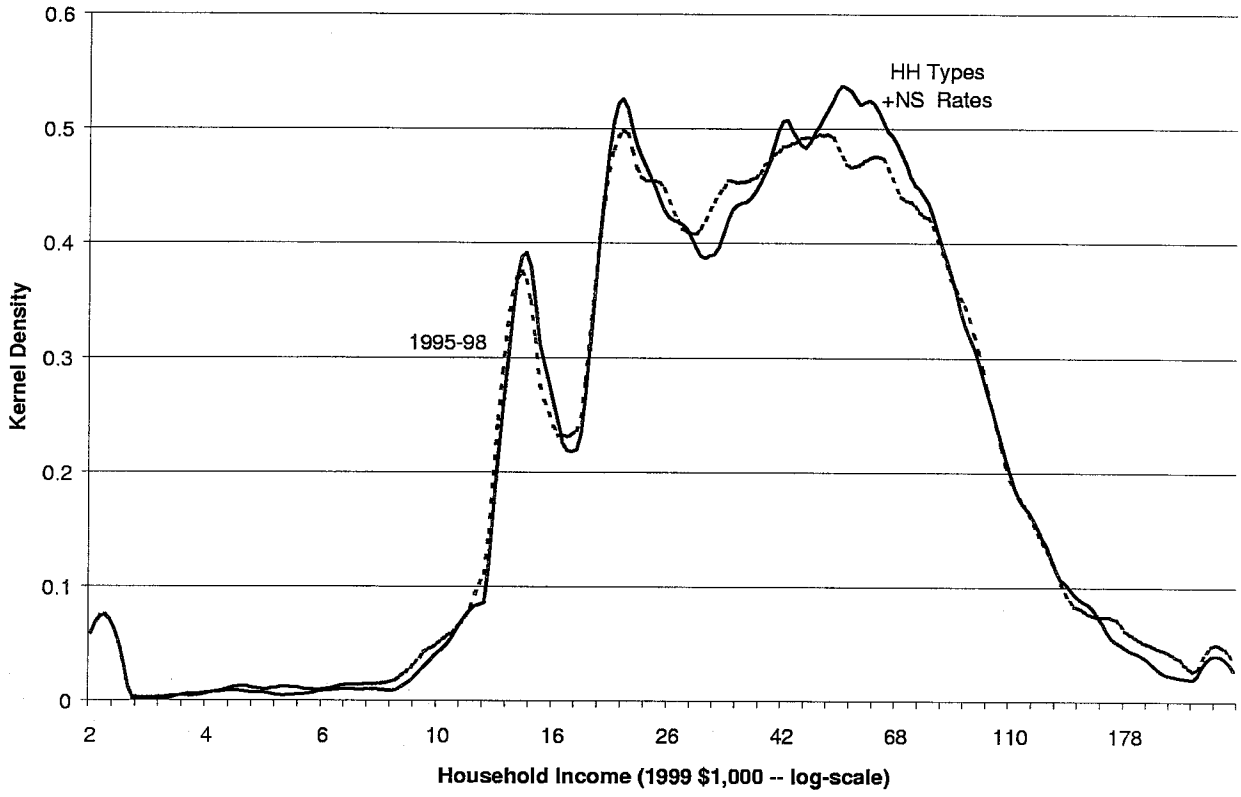

Figure 14b: Changes Between 1989-92 and 1995-98: National Superannuation Contribution