The Australian National University
Centre for Economic Policy Research
DISCUSSION PAPER

The Effects of Assortative Mating on Income Inequality: A Decompositional Analysis

Shane Mathew Worner

DISCUSSION PAPER NO. 538
November 2006

ISSN: 1442-8636
ISBN: 0 921262 09 5

Currently a PhD scholar in the Research School of Social Sciences (RSSS), The Australian National University. Address: Economics Program, Research School of Social Sciences, Coombs Building 9, The Australian National University, Fellows Road, Acton ACT 0200.
shane.worner@anu.edu.au

Acknowledgements
The author would like to thank Juan David Baron, Bob Breunig, Bruce Chapman, Andrew Leigh and participants at the RSSS Seminar series for helpful ideas, advice and comments.

The data used in this study are from the Survey of Income and Housing Costs (SIHC). SIHC is a confidentialised file arranged and managed by the Australian Bureau of Statistics (ABS). The findings and views expressed in this paper are that of the author and should not be attributed to the ABS.

ABSTRACT

Using the Australian Bureau of Statistics' Survey of Income and Household Costs, this paper explores the effect of changing assortative mating patterns on income inequality. Evidence from theoretical and mathematically calibrated models suggest that assortative mating has distributional implications for measurable traits, which include income. Using a semi-parametric conditional weighted kernel density estimation framework we analyse the effect of assortative mating on the distribution of income in Australia. In controlling for labour force participation, family characteristics, education and other demographic variables, we find some evidence to suggest that assortative mating has had an influence on the increase in income inequality in the 17 years to 2003. The results are robust to several changes in specification.

JEL Classifications: D31, D63

Keywords: inequality, assortative mating/matching, sorting

1. Introduction

Over recent years there has been a number of works investigating the distribution of wages in Australia. Most agree that there has been a dramatic increase in the degree of income inequality in the 1980's and early 1990's, whether it is at the household, income unit or individual level (Borland and Wilkins 1996; Barrett et al 2000; Blacklow and Ray 2000; Harding and Greenwell 2002; Johnson and Wilkins 2004; Leigh 2005). The findings suggest that the widening of the distributions in Australian incomes has many driving forces. Changes in the distribution of employment, declines in income at lower ends of the distribution followed by increases at the upper levels, changes to marginal tax rates especially at the higher end, other Government efforts to redistribute income and increased female labour force participation are just several reasons presented in the current literature to help explain the changing trends in income distributions over the past few decades. The results are robust, regardless of whether income or expenditure is chosen for the inequality analysis. However, to date there is one hypothesis that has yet to be tested, that is, how much of the increase in income inequality over the past two decades can be explained by changes in assortative mating patterns and how quantitatively significant are these channels? These are important questions that this paper intends to address. Assortative mating is defined as the partnering of individuals with more traits in common than would be likely through random partnering. Random partnering would suggest no partnering pattern in the traits brought to a partnership, which empirically is not the case (Mare 1991).

This paper investigates the extent to which changed assortative mating patterns contributed to the change in the Australian income distribution between the years 1986 and 2003.[1] Using household unit level data provided by the Australian Bureau of Statistics' (ABS) Survey of Income and Household Costs (SIHC) for 1986 and 2003, the question is examined using a modified version of an increasingly popular decomposition method proposed by DiNardo, Fortin and Lemeuix (1996) (DFL from this point on).

[1] For the purpose of our analysis we do not restrict the definition of mating to those individuals that are legally married.

This paper adds to the current literature on this specific topic in several ways. First, as part of the preliminary analysis we investigate the extent of assortative mating in Australia. To the best of the author's knowledge this has yet to be investigated to date. Second, very little empirical work has focused on the effects of assortative mating on income inequality, though theoretical and some calibrated mathematical models suggest that there should be a significant effect. Third, our analysis provides a robust causal framework to determine if and by how much changed assortative mating patterns have influenced income inequality over time.

The rest of the paper is set out as follows. Section 2 presents some of the theoretical and empirical implications of assortative mating with Section 3 introducing the data and summary statistics. In addition, Section 3 also provides a brief commentary on measures of assortative mating and income inequality used in the analysis. Following this, Section 4 outlines the estimation methodology with Section 5 presenting the results. Section 6 tests the robustness of the results to differing definitions of equivalised income and assortative mating. Section 7 concludes.

In the first chapter we investigated the educational effects on the timing of marriage. This current paper follows on from this in several ways. First, the timing of marriage (and more importantly family formation) can truncate the accumulation of human capital. This can potentially lead to less favourable labour market positioning and therefore can impact on inequality in income between households. Second, the timing of marriage can have important implications for educational assortative mating. The age at marriage along with educational attainment has been increasing. With much of the increase in educational attainment since the 1960's being a result of increased female participation, this implies that spousal resemblance on educational qualifications is likely to become more similar. Once again, if educational qualifications translate directly into wages and salaries, increased sorting on spousal educational traits will impact on the distribution of household incomes.

2. Conceptual and Empirical Issues

It has been argued that increased assortative mating can have important implications for the distributions of several social and economics factors, including wages and income. Economic models suggest that increased sorting leads to increased inequality in the income distribution, although of recent times controversies have arisen to the extent of its true effect in simulated results. Becker (1973; 1991) points out that not only does assortative mating decrease the intergenerational variance of traits within families but in particular positive assortative mating also increases the inequality between families. However, his theory also predicts that the effect on wage income within the household will be quite different from that of the commodity output of a partnered couple. If wages rates are negatively sorted (although not empirically supported (Lam 1988)), as suggested by Becker, assortative mating should reduce inequality in money income between families. This implies that if wages were positively sorted it would increase inequality. So regardless of whether the trait of interest is a productivity measure (wages) or income measure (commodity output), positive assortative mating will increase the dispersion in that trait.

Burdett and Coles (1997) find that within a matching framework that consists of heterogeneous agents and non-transferable utility, individuals with similar traits will segment themselves into different groups within the marriage market. Essentially they form different "classes", which are defined as a function of the individual's respective traits. A result of the market equilibrium is that an individual will only partner with a person within the same class. A male from a high class will not partner with a female from a lower class and vice versa, implying positive assortative mating among individuals. An interesting implication of this, as pointed out by the authors, is that if assortatively mated couples were to have children, who inherit a weighted average of the traits that the parents brought to the partnership, assortative mating of individuals can have intergenerational effects on the distribution of those traits brought to the marriage, making them more unequal overtime. Fernandez *et al* (2001) find a similar result. Increases in sorting are likely to have a quantitatively significant effect on the level of income inequality. Their results hold even in the absence of borrowing constraints, with market imperfections likely to contribute to a magnification of inequality.

The controversy arises, however, with Kremer (1997) suggesting that sorting has little impact on inequality of characteristics that are only moderately inheritable such as education and income. Kremer observes that the degree of sorting in the U.S, as measured by the correlation between spouses' education levels, has changed very little over the fifty years to 1990. Given this, simulated results suggest that a dramatic increase in spousal mating would only increase the steady state deviation of education by 1 percent, with a doubling in sorting patterns increasing the standard deviation by 1.7 percent. The author points outs that this change in the steady state standard deviation is the equivalent of an extra 6 days worth of education a year. Although the focus is on education inequality the implications for income inequality would be similar, as the author points out since "… the intergenerational correlation of income is approximately 0.4…" in a US context.[2]

Empirically, there has been little done in the field of income inequality and assortative mating. Of the empirical work that has been done in this area, the results are mixed. At a multi country level, including Australia, Fernandez *et al* (2001) investigates the relationship between marital sorting along educational lines and differing measures of skills premia, which include the male female wage ratio, a skills indicator and a Mincer coefficient. Their findings indicate that for a number of countries, Australia inclusive, there is a significant positive relationship between marital sorting (as measured by the Pearson correlation coefficient of partner's education) and the varying measures of skills premia. However, due to the OLS techniques implemented in the analysis one criticism of the results is that causality is indeterminable, only providing a slightly more robust measure of correlation. In addition, income inequality can and does effect the entire distribution. It can have quite large and distinct effects at either "tails" of the distribution. Although the information obtained from the OLS regression analysis may be anecdotally useful for cross-country comparisons, it only provides clarification at one specific point, which is at the mean. It therefore falls well short of providing insight into how sorting patterns affect the entire distribution of income. The econometric decomposition method employed in this paper addresses this concern and is discussed further in Section 4.

[2] Marks *et al* (2005) show that for Australian households there is little correlation between household wealth and parental characteristics. Although wealth, as pointed out by the authors, is much more unequal, it is, however, strongly related to an individual's income and provides some anecdotal evidence that a person's wealth or income is not highly related to some intergenerational characteristics.

Tsai and Kuan (2004), using a quantile regression framework within a single cross section, investigate the effects of both occupational and educational homogamy on the income distribution of Taiwanese families.[3] Their findings indicate two things; one, educational homogamy has a positive effect on family income and two, educational homogamy has no significant effect on any inter-quantile relationships and subsequently provides very little evidence to support the hypothesis that increased educational sorting leads to a significant widening of the family income distribution. Moreover, they find that occupational sorting is more likely to have a greater polarising effect on income and its distribution than educational matching.

3. Data and Descriptive Analysis

The data used in the analysis are from the household unit record files of the SIHC conducted by the ABS in the years 1986 and 2003. After accounting for data irregularities, the combination of the two SIHC surveys contains 33,468 individual observations in total and includes information on age, sex, martial status, labour market outcomes and labour force participation (including occupation, hours worked, industry of occupation), education qualification details and information on family characteristics and demographics.[4] Table 1a and Table 1b presents the summary statistics of the variables considered in the econometric model for each year in the analysis.

The household unit is the basic unit for which the inequality analysis takes place and we restrict our analysis to those individuals that are aged between 16 and 64.[5] No regions of Australia have been excluded from the analysis. Since our focus is on how

[3] Educational Homogamy refers to people with like educational qualifications partnering with each other. It is analogous to positive assortative mating on education.

[4] The individual number of observations for each year is as follows: 1986: 14,990; 2003: 18,478. These numbers are after accounting for data irregularities in the samples, such as non-reporting of education and marital status.

[5] In limiting the sample to those that are working age we exclude those individuals that are retired and rely upon government funded pensions. This restriction is common in the current literature (see Hyslop and Mare 2005; Daly and Valletta 2006) and allows for comparison of our work, although Fernandez et al (2001) suggest that younger cohorts (i.e. those aged between 16-25) should not be included because they are presumably less stable in regards to their partnering patterns.

changes in assortative mating patterns have impacted on the distribution of household incomes in Australian households, the income measure of choice is the household's

Table 1a: Summary Statisitics and Variable Descriptions

Year			**1982**		
Variable	Obs	Mean	Std. Dev.	Min	Max
OECD Equiv Weekly Inc.	14990	234.1355	197.5739	0	3172.353
Per Capita Equiv Weekly Inc.	14990	204.3342	174.2251	0	2696.5
Age	14990	37.7521	13.78614	16	63
Highest Education	14990	2.514009	1.493521	1	5
Group: Labour Force Participation					
Fulltime	14990	0.5329553	0.4989294	0	1
Partime	14990	0.1392929	0.3462634	0	1
Unemployed	14990	0.0611074	0.2395352	0	1
Not In Workforce	14990	0.2666444	0.4422197	0	1
Hour worker per Week	14990	25.63696	20.44403	0	50
Group: Family Charateristics					
No of Dependant kids	14990	0.9104736	1.202435	0	6
Group: Assortative Mating					
Dichotomous Variable	14990	0.4838559	0.499756	0	1
Gender median Distance def.	14990	0.8877366	2.779657	-4.339687	18.62013
Group: Education & Demographics					
Born Overseas	14990	0.2622415	0.4398679	0	1
Female	14990	0.5016678	0.5000139	0	1
NSW	14990	0.2470981	0.4313386	0	1
Victoria	14990	0.206471	0.4047859	0	1
Queensland	14990	0.169513	0.3752169	0	1
South Australia	14990	0.1356905	0.342471	0	1
West Australia	14990	0.139026	0.3459852	0	1
Tasmania	14990	0.0670447	0.2501077	0	1
ACT/NT	14990	0.0351568	0.1841821	0	1
Yr10 or Below	14990	0.4557705	0.4980565	0	1
Yr11 or Equiv	14990	0.0321548	0.1764169	0	1
Yr12 or Equiv	14990	0.1312208	0.337653	0	1
Post School Cert/Dip	14990	0.3040027	0.4599991	0	1
Bachelors and Other Post Grad	14990	0.0768512	0.2663641	0	1

Notes: 1. Summary statistics supplied for Income and hours worked represent weekly totals.
2. Groups and variables listed are the ones used in the decompostion with some interaction variables ommited from the table 3. Summary Statistics presented are for the working age population aged between 16 and 64

Source: Australian Bureau of Statistics' Survey of Income and Housing Costs 1986

Table 1b: Summary Statisitics and Variable Descriptions

Year		2003			
Variable	Obs	Mean	Std. Dev.	Min	Max
OECD Equiv Weekly Inc.	18478	551.3546	478.5998	0	6335.294
Per Capita Equiv Weekly Inc.	18478	475.5841	413.988	0	5385
Age	18478	39.65943	13.38275	16	64
Highest Education	18478	3.110023	1.476766	1	5
Group: Labour Force Participation					
Fulltime	18478	0.505953	0.4999781	0	1
Partime	18478	0.2216149	0.4153445	0	1
Unemployed	18478	0.0425371	0.2018164	0	1
Not In Workforce	18478	0.229895	0.4207765	0	1
Hour worker per Week	18478	26.28523	19.55932	0	50
Group: Family Charateristics					
No of Dependant kids	18478	0.862431	1.13528	0	5
Group: Assortative Mating					
Dichotomous Variable	18478	0.553036	0.4971927	0	1
Gender median Distance def.	18478	0.488514	1.881373	-5.366784	9.078008
Group: Education & Demographics					
Born Overseas	18478	0.2510553	0.433632	0	1
Female	18478	0.5139626	0.4998185	0	1
NSW	18478	0.2322221	0.4222613	0	1
Victoria	18478	0.2135513	0.4098246	0	1
Queensland	18478	0.1728001	0.3780846	0	1
South Australia	18478	0.1065592	0.3085603	0	1
West Australia	18478	0.1284771	0.3346293	0	1
Tasmania	18478	0.0681892	0.2520771	0	1
ACT/NT	18478	0.0782011	0.2684951	0	1
Yr10 or Below	18478	0.2598766	0.4385786	0	1
Yr11 or Equiv	18478	0.0745752	0.2627117	0	1
Yr12 or Equiv	18478	0.1483927	0.3554984	0	1
Post School Cert/Dip	18478	0.32996	0.470211	0	1
Bachelors and Other Post Grad	18478	0.1871956	0.390079	0	1

Notes: 1. Summary statistics supplied for Income and hours worked represent weekly totals.
2. Groups and variables listed are the ones used in the decompostion with some interaction variables ommited from the table 3. Summary Statistics presented are for the working age population aged between 16 and 64

Source: Australian Bureau of Statistics' Survey of Income and Housing Costs 2003 gross weekly income from wages and salaries.[6] Our analysis focuses on the log of this variable. Household income is adjusted to account for the household size using the OECD equivalence scale, which assigns a value of one for the first adult in the income unit with every subsequent adult being assigned a value of 0.7. Each child in the income unit is assigned the value of 0.5.

The OECD equivalence scale is a standard scale that is used in OECD reporting. [7] Other standard equivalence scales used in the literature include Engel equivalisation, demand system based scales (DSBS) and per capita equivalisation. For the purposes of this paper, we have chosen the OECD equivalence scale for the primary decomposition. The major advantage of using Engel and DSBS equivalence scales is that they take into account economies of scale within the family or household unit. However, they were not introduced into this decomposition because Engel and DSBS equivalence scales not only require substantial computational effort but they also require price and expenditure information on a basket of goods that make up the family or household budget. The OECD equivalence scale provides us with a simple way of account for the number of people within the household, and since the weightings decrease with the type of individual in the household unit, it also implicitly captures some sort economies of scale effect in relation to the size of the household unit. The robustness of the results will be tested to the OECD equivalence scale further in the paper using several other standard equivalisation methods as an alternative.

Formally, if there are n number of individuals who are members of m number of households and Y is our measure of income for household j, then:

$$Y_j = \frac{Y_{HUj}}{OCEDEquvi_j} \qquad j = 1,...,m \tag{1}$$

[6] Normally inequality analysis utilises only that income that is available to the household, namely disposable income. As Harding (1997) points out, however, Government taxes and transfer payments, on average, fully offset any market based increases in income inequality in Australia between the years 1982-1994. Therefore to help untangle the consequences of assortative mating, free of Government redistribution mechanisms, gross income from wages and salaries is used.

[7] The primary decomposition uses the "old" OECD equivalence scale (sometimes known as the Oxford scale) where the first adult house member is assigned the value of 1 with all subsequent adults assigned a value of 0.7. Each child is assigned the value of 0.5. The robustness of the results is tested in Section 6 to differing equivalisation scales. For a further discussion of Equivalence Scales consult Lancaster and Ray (1998) and Blacklow and Ray (2000).

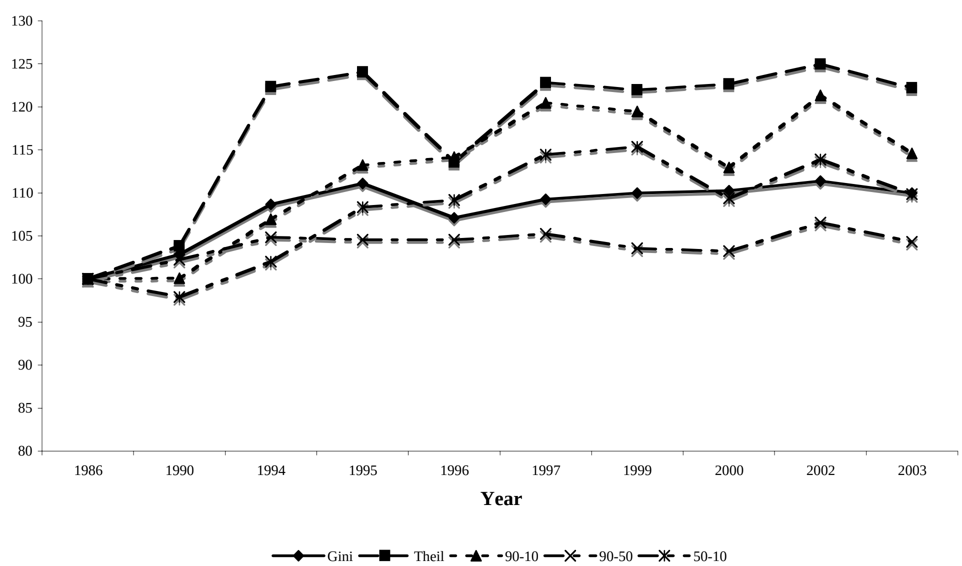
where $Y_{HUj} = \sum_{i=1}^{n} Y_{ij}$ is weekly household unit income, which is defined as the summation of all incomes of individuals $1,...,n$ who are a member of household j. $OECDEquvi_j$ is the calculated OECD equivalence scale for household j.

Finally, all household incomes are deflated using the Consumer Price Index (CPI) as reported by the ABS and are expressed in 2006-dollar terms.[8]

3.1. Trends in Income Inequality and Assortative Mating

As part of the preliminary analysis we examine how both income inequality and assortative mating patterns have changed in Australia from 1986 to 2003. Figure 1 presents the relative change in a number of income inequality measures.[9] Figure 2 presents the trends in incomes in Australia between the years 1986-2003. Each of the inequality series is indexed to 100 with 1982 being the base year of comparison, with the income indicators presented in 2006 dollars.

Figure 1: Relative Changes in Income Inequality 1986-200



[8] In the two SIHC samples we have excluded from the analysis those observations whose log real incomes were 0 (i.e. $1 per week). The data is not right censored, that is the data is not top coded.

The inequality measures presented in the graph are the standard in the inequality literature and include GINI, Theil, 90-10, 90-50 and 50-10. GINI is the most common measure on inequality employed and is described as the ratio of the area between the equality line and the Lorenz curve to the total area under the equality line.[10] It takes a number between 0 and 1 with 0 indicating perfect equality and 1 being perfect inequality. The two major criticism of the GINI co-efficient is that it, one, only looks at inequality amongst one group and, two, is sensitive to movements in income in the middle of the distribution than at the extremes of the distribution.

The Theil co-efficient is another common measure of income inequality that falls into the entropy class of inequality measurements. Similar to the GINI co-efficient, the Theil co-efficient takes a value between 0 and 1. However, it has one major advantage over the GINI co-efficient. It not only measures the inequality in the distribution across groups, but it also takes into account the distribution of the individuals within those different groups. Therefore groups that have an equal share of the (income) distribution contribute nothing to the overall Theil co-efficient. Hence, if all groups have and equal share of income, the Thiel co-efficient will be zero and is interpreted as perfect equality.[11]

90-10, 90-50 and 50-10 are described as dispersion ratios, and as their description implies, they are a ratio of their respective deciles. For example, 90-10 is simple a ratio of income at the 90th deciles and income at the 10th decile. This gives us an intuitively simple indication of how many times more people at the upper end of the distribution earn compared to those individuals or households at the lower end of the distribution. Similarly 90-50 is a ratio of the 90th decile to the median income with 50-10 being a ration of the median income to the 10th decile.

Inspection of both graphs reveals some interesting trends. First turning our attention to inequality, of the inequality measures presented, all measures increased during the time period 1986 – 2003, with the majority of the increase in income inequality

---

[9] All Income unit income inequality measures are calculated on OECD equivalised income. As a robustness check the same inequality measures were calculated using Per capita equivalised income. The results are presented in the Appendix.

[10] See Appendix for a graphical representation.

[11] For a further discussion on the Theil measure of inequality, please consult Conceição and Ferreira (2000)

coming early in the time period. The trend in equality over the period is most evident in the bottom half of the income distribution. As represented by the 50-10 percentile ratio and the Theil measurement of inequality, which is weighted more heavily towards the bottom end of the distribution, the graph shows a 8 percent and 24 percent increase in both measures respectively from 1986-1995. The other inequality measures show a similar trend, with varying degrees. The 90-10 dispersion ratio shows a large percentage increase in inequality, with the ratio increasing 20 percent to 1997. When viewed in conjunction with the 90-50 dispersion ratio, which has remained relatively constant in the 17 years to 2003, much of the increase in inequality is due to the movement in the top half of the income distribution.

Since 1995-1996 there have been a “flattening out” of the various measures of income inequality, with the Theil measurement showing the greatest increase since 1996. There has been only minimal increase reported in the other measures to 2003. These trends in gross income inequality are consistent with similar results from several authors (Harding and Greenwall 2002; Johnson and Wilkins 2004; Leigh 2005).

Figure 2: OECD Equivalised Income Indicators (1986-2

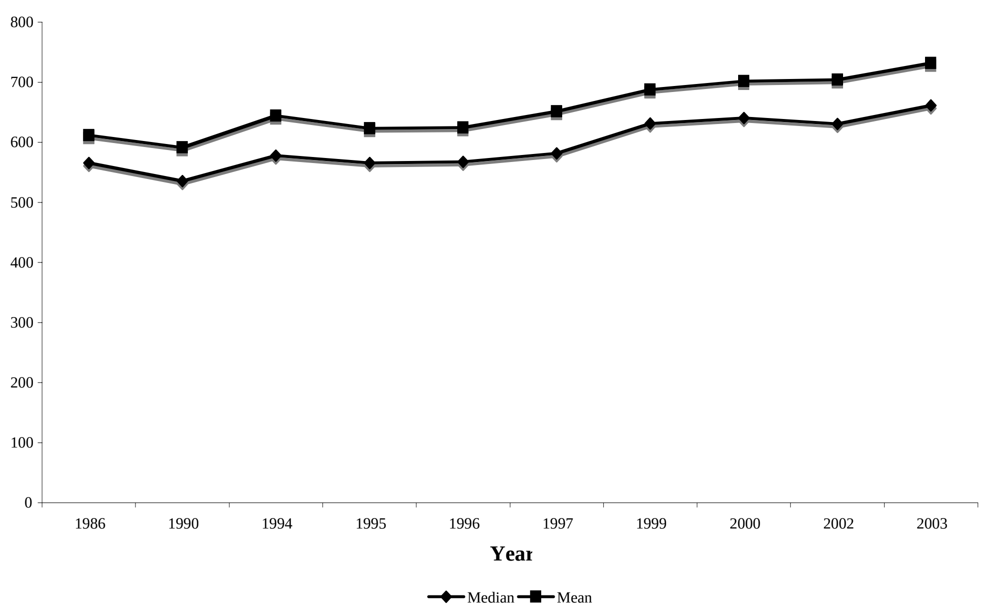


Figure 2 graphically presents trends in income in Australia over the 17 year period. Between 1986 and 2003 the median and mean equivalised incomes increased over the time period in question. Both measures of income show a 5 percent decline in the late 1980's but since have increased consistently through to 2003 with both the median and mean income levels increasing 16 percent and 19 percent respectively to 2003.

3.1.1. Simply Density Estimates

We next turn our attention to the changes in the income distribution in Australia during the period 1986 -2003. The densities are estimated using a weighted kernel density function, which is quite a common way of graphically representing the distribution of a continuous, measurable trait. The weighted kernel density function takes the form

$$\hat{f}_t(y) = \sum_{i=1}^{N} \frac{\theta_{ti}}{h} . K\left(\frac{y - y_{ti}}{h}\right) \qquad (2)$$

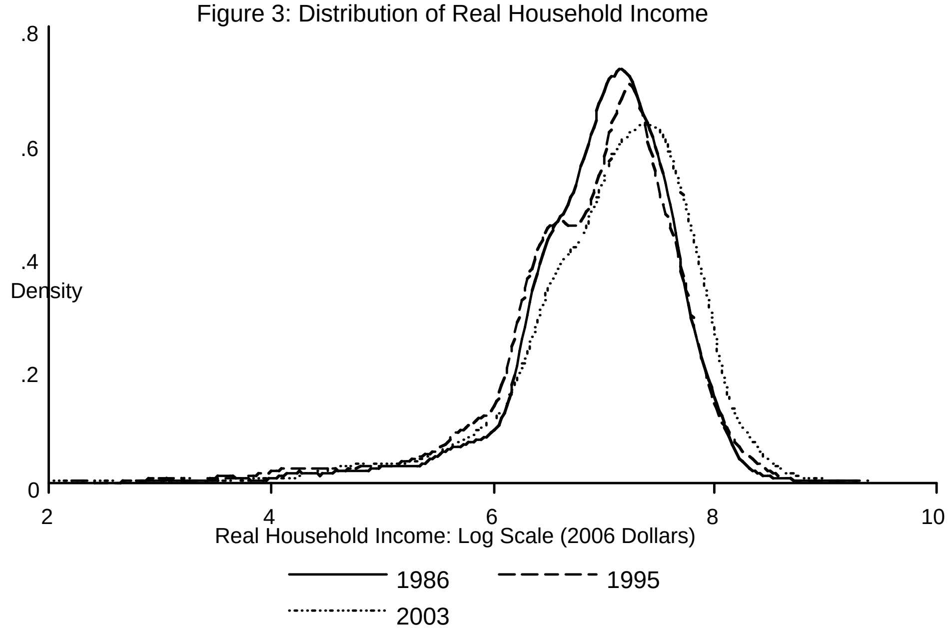
where h represents the kernel bandwidth and K is the Kernel function.[12]

Figure 3 provides a kernel density estimation of real household unit incomes for 1986, 1995 and 2003. Log household unit income has been converted into 2006 dollars.[13] Inspection of the graph reveals that the distribution of incomes for both 1986 and 2003 follows, for all intents and purposes, a somewhat bell shape pattern. The figure also reveals that there has been a widening in the distribution between the years represented. Between 1986 and 1995 there has been a hollowing out of the middle of the distribution with the tails at each end of the income distribution becoming "fatter". In addition, between 1986 and 2003 there has been a real shift in the distribution of income with the mass of the distribution being right shifted. This represents a real increase in the real incomes of Australian households and is anecdotally confirmed by the real increase in the mean and median incomes presented in Figure 2.

[12] For the purposes of this paper, the densities are estimated using an Epanechnikov kernel of the form $K(z) = 0.75[(1 - z^2)/5]/\sqrt{5}$, if $|z| =< \sqrt{5}$: and $K(z) = 0$ otherwise. Bandwidths have been set to the "optimal" width as calculated by STATA. STATA explains that the optimal bandwidth is "… that bandwidth that minimises the mean integrated square error…"

[13] Incomes are deflated using the Consumer Price Index as advised by the ABS.

Figure 3: Distribution of Real Household Income



3.1.2. Measurements of Assortative Mating

We now turn our attention to the extent of assortative mating amongst Australian couples. The variable of choice used to determine the extent of assortative mating is an individual's education level. Why? Education can have significant effects on the productivity of individuals in both the household and labour markets sectors and therefore may have and influence on the decision to partner and on the decision of who partners with whom. Also, education levels are good predictors of labour market outcomes but, in addition, also allows for comparison of the following preliminary work on assortative mating with the current empirical literature in this area. The education variable in the two SIHC surveys is a 5 level ordinally ranked variable that takes the value of 1-5.[14]

Focusing now on the preliminary analysis of the data, one of the most elementary measurements of assortative mating is a cross tabulation, between partners, of the traits of interest. For the purposes of this preliminary analysis, "partners" is defined as those individuals who are legally married. Table 2 presents a cross tabulation of an individual's education level to that of their partners for each year of the study. The most rudimentary measure of assortative mating is simply the aggregation of the

[14] Education levels are defined as the following: 1. Year 10 or Below; 2. Year 11 or Equiv; 3. Year 12 or Equiv; 4. Post School Certificate/ Diploma 5.. Bachelor's Degree and other Post Graduate

diagonal elements. We find that between 44 percent and 42 percent of working age individuals are on the main diagonal sum of the table across the two sample years. This result is in line with similar findings in the assortative mating empirical literature. For example, Halpin and Chan (2002) found that 49.3 percent and 39.6 percent of Irish and English observations respectively are on the main diagonal. Similarly Çelikaksoy et al (2003) finds a similar figure amongst native Danes at 45 percent.

The major criticism of this measure of assortative mating, however, is that as education levels become more scattered across different schooling levels, reported assortative mating levels would tend to decline. Halpin and Chan (2002) point out that this measure of assortative mating is quite sensitive to marginal distributions in education. That is, this particular measure does not effectively hold constant changes in the marginal distributions of educational attainment. In fact levels of assortative mating are quite closely linked to the dispersion of education levels. The more uneven is the distribution of educational levels the more likely we will see large designations of positive assortative mating even if "… education has no bearing on partnering decisions".

Therefore, in a more concrete analysis of assortative mating, we use Kendall's Tau to measure the correlation between an individual's level of educational attainment and the education level of their partner. Kendall's Tau measures the extent of correlation between two ordinally ranked variables. It is the appropriate measure since it exhibits small sample properties.[15] Kendall's Tau requires that the two variables, X and Y, be paired in some way. An example of this would be an individual's education level and the education level of their partner. As long as the variables are ordinally ranked, then the correlation between them can be measured. Kendall's Tau is carried out on the rank of the data. So for each variable the values are placed in order of rank, 1 for the lowest; 2 for the second lowest and so on. As with Spearman's Rho, Kendall's Tau takes a value between -1 and +1 and has an intuitively simple interpretation of the

[15] Kendall's' Tau is defined as $\frac{S}{N}$ where S is a score (Kendall's score reported in Table 3) defined by C-D. C and D are the number of concordant and discordant pairs respectively. N is the total number of pairs and is define as $N=\frac{n(n-1)}{2}$.

Table 2: Crosstabulations of Educational Qualifications

1986

Highest Qualifications	Spouses Education: Yr10	Yr11	Yr12	Post School Certificate	Bachelors and other PG
Yr10	958	31	116	527	34
Yr11	31	8	10	33	7
Yr12	115	10	56	142	28
Post School Certificate	527	33	142	475	93
Bachelors and other PG	34	7	28	93	122
Total	1665	89	352	1270	284
					3660
Diagonal					1,619
Diagonal Percentage (%)					**44%**

2003

Highest Qualifications	Spouses Education: Yr10	Yr11	Yr12	Post School Certificate	Bachelors and other PG
Yr10	1,307	155	241	984	151
Yr11	155	76	86	266	81
Yr12	243	87	306	491	256
Post School Certificate	984	269	493	1,794	611
Bachelors and other PG	147	81	255	611	1,294
Total	2,836	668	1,381	4,146	2,393
					11,424
Diagonal					4,777
Diagonal Percentage (%)					**42%**

Source: Australian Bureau of Statistics' Survey of Income and Housing Costs 1986-2003 strength of the relationship between two variables. A positive number indicates that the rank of both variables, X and Y, increases together. A negative number indicates that as the rank of one variable increases, the rank of the other variable decreases. Another way of interpreting Kendall’s tau is as the difference in probabilities, or more concretely, the difference between the probability of concordance and the probability of discordance.[16] For example, say we have education levels for two separate groups of individuals, and further, the Kendall’s Tau correlation co-efficient on education between the two groups is 0.30. This co-efficient means that if we were to take a person from each group, there is a 30% chance that their education levels will be similar rather than dissimilar. The other advantage of Kendall’s tau is that it allows for hypothesis testing between the variables with the null hypothesis being that the variables are independent.

Table 3: Kendall's Tau Correlation Measures for 1986-2003

Year	**1986**
No. Obs	3660
Kendall's tau	**0.1511*** **
Kendall's score	1011962.00
	(63682.02)
Year	**2003**
No. Obs	11424
Kendall's tau	**0.2352*** **
Kendall's score	15348237.00
	(376717.84)

Note: 1. Kendall's Tau Correlation measure calculates the correlation between an individual's education level and the education level of the spouse. 2. Standard Errors of Kendall's score are in parenthesis 3. *** indicates co-efficent is significant at 1% level,

Source: Australian Bureau of Statistics' Survey of Income and Housing costs 1986-2003

Inspection of Kendall’s Tau scores presented in Table 3 reveal that the within correlation for each year is significantly positive. When interpreting the strength of the relationship, there is no consensus as to what constitutes a weak, moderate or

[16] Concordance is defined as the degree of similarity in a pair of individuals with respect to a given measurable trait. Conversely, discordance is defined as the degree of dissimilarity in a pair of individuals with respect to a given measurable trait. More formally with respect to Kendall’s tau, if there are two sets of observations $(X_i, Y_i), (X_j, Y_j)$. If the product of $(X_j - X_i)$ and $(Y_j - Y_i)$ is positive, then they are said to be concordant pairs. Conversely, if the product of $(X_j - X_i)$ and $(Y_j - Y_i)$ is negative, then they are said to be discordant pairs.

strong relationship. However as a general rule of thumb: weak relationship, ±0.01 to ±0.30; moderate relationship, ±0.31 to ±0.70; strong relationship ±0.71 to ±0.99. A perfect relationship is ±1.00 and no relationship is 0.00.[17] Given these guidelines, we conclude that the relationship between partner's education levels is weak but has increased between the two samples. Although the magnitude of the correlation coefficients suggests that the relationship is weak, for both SIHC samples used in the analysis, the results confirm that some positive assortative mating exists between partners' education levels.[18]

To test the difference in magnitude between the two independent correlations, we use a Fischer Z transformation to determine whether one correlation is significantly different from the other.[19] One of the statistical properties of the Kendall's Tau statistic is that it is normally distributed, allowing for z-score statistical inference testing. However, the main assumptions underpinning normality is that the sample size be significantly large, which we have, and that there is not too many ties in the ranking of the data, which we do not have.[20] This is the reason why we undertake a Fischer transformation to correct for non-normality.

In testing the correlation coefficients of 1986 and 2003, we find a significant difference between the coefficients. The results are presented below in Table 4. This suggests that compared to 1986 the correlation of partners' education levels is statistically significant and, therefore we can conclude that the level of assortative mating by education, as given by this definition, has significantly increased in the 17 years to 2003.

---

[17] See Elifson (1982)

[18] The result is consistent with correlation results found by Fernandez et al (2001). Although the magnitude of their results is much larger, their analysis uses Spearman's Rho as the correlation measure, which tends to be larger in absolute value than Kendall's Tau.

[19] Fischer's Z transformation is defined as $Z' = \frac{1}{2}(\ln(1+\tau) - \ln(1-\tau))$. The difference between the two transformed correlations is then divided by the standard error, which is given by $\sigma_1 - \sigma_2 = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$. If the score is greater than 1.96 then we conclude that the difference between the two independent correlations is significant at the 5% level.

[20] For further discussion on the asymptotic properties of the Kendall's Tau statistic, please refer to Hays (1988) pg 840-844

Table 4: Fischers Transformation of correlation coefficents for 2003 and 1986

Year	**2003**
No. Obs	11424
Kendall's tau	0.2352
Fischers Z_1	0.2386

Year	**1986**
No. Obs	3660
Kendall's tau	0.1511
Fischers Z_2	0.1522

Difference between 2003 & 1986	
Z_1-Z_2	0.0864
SE	0.0003
Fischers Score	**288*** **

Note: 1. Fischers Score measure the significance of the differnce in correlation between 2003 & 1986
2. *** indicates co-efficent is significant at 1% level,
Source: Australian Bureau of Statistics' Survey of Income and Housing costs 1986-2003

For the purposes of the analytical section of the paper, we introduce two different definitions of assortative mating into the model. The main definition used in the primary decomposition is defined as a dichotomous variable. Since economic theory in this area postulates that positive assortative mating (i.e. "likes" partnering with "likes") will increase inequality in measurable traits, the binary variable will denote positive assortative mating (PAM) or otherwise.[21] PAM is defined as having an education that is above (below) the gender specific median and also partnering with an individual who has an education level above (below) their gender specific median.[22] For example, if a male or female with a university education partnered with a male or female who was also educated to a university level, these individuals would be considered PAM on education. Similarly, if a male or female with schooling only to a year 10 standard partnered with a male or female with year 10 qualifications, they too would be considered PAM on education.[23] Table 1 indicates that between 48

[21] Otherwise captures all those who are negatively assortatively mated or single and therefore assumes that the two groups are similar. This assumption is relaxed in robustness checks in section 5 of the paper

[22] We use gender specific median in this case to account for the fast that the distribution of education differs between females and males

[23] For the purposes of determining which individuals are PAM or otherwise, the specific education medians for each gender and each year are: Male: 1986 = Year 12, 2003 = Post School Cert. Female: 1986 = Year 10, 2003 = Year 12

percent and 55 percent of individuals in the analysis are PAM, as defined by this definition, for the years 1986 and 2003 respectively and indicates that assortative mating, on average, has increased over this time frame.

The advantage of this particular definition of assortative mating is that is provides an intuitively simple way of determining whether an individual exhibits PAM or not. However, there are two major drawbacks in using this measure. First, by turning the assortative mating variable into a dichotomous variable, we lose a lot of information on the strength on the mating pattern. That is, we cannot determine how assortatively mated an individual is. Second, this measure is susceptible to the number of categories in the underlying variable. The larger the size of the bands within the underlying variable or the smaller the number of categories, the larger the number of people who are captured by the category and therefore the larger the reported extent of assortative mating between partners. To account for the fact that this measure of assortative mating is susceptible to the banding within the determining variable we introduce a second definition of assortative mating.

The second measure of assortative mating is defined as the product of the two partners' distances from their gender respective medians, weighted by their standard deviations to account for the difference in the distribution of education between males and females. It is presented in conjunction with our first definition of assortative mating as an alternative in the primary decomposition.

Formally, if there are p number of partnerships, then the degree of assortative mating of individual i, who is member of couple k at time t, is determined by the following equation:

$$ASM_{ik,t} = (\tilde{f}_{ik,t} - f_{median,t})(\tilde{m}_{ik,t} - m_{median,t}) \qquad (4)$$

$$t = 1986, 2003$$
$$k = 1, ..., p$$

where f_{median} and m_{median} are the gender specific education medians at time t, $\tilde{f}$ and $\tilde{m}$ are defined as the male and female education levels at time t, weighted by their respective sample standard deviations.

Therefore,

$$\tilde{f}_{i,t} = \frac{\hat{f}_{i,t}}{s_{f,t}} \quad \& \quad \tilde{m}_{i,t} = \frac{\hat{m}_{i,t}}{s_{m,t}} \qquad (5)$$

Not only does this definition of assortative mating provide a continuous variable that accounts for all degrees of assortative mating with a negative number indicating negative assortative mating (NAM) and a positive number interpreted as evidence of PAM. In addition, it also allows for those individuals that are not partnered by assigning a value of 0. The summary statistics in Table 1 indicate that, on average, partnerships exhibit PAM across the education variable for the 1986 and 2003 samples respectively, with the magnitude increasing across the two years.

4. Estimation Methodology

As stated previously, the estimation technique employed by this paper is analogous to that proposed by DiNardo, Fortin and Lemeiux (1996). It is a semi parametric distributional framework that has several advantages. First, it provides a quantitative or qualitative mechanism by which to view the distributional effects of chosen factors. Second, it provides greater insight into the household income distribution at a specific point in time. Third, the DFL method works with the entire density of the income distribution, unlike another popular decomposition method, the Oaxaca decomposition. The Oaxaca decomposition (1973) allows for the breakdown of the difference in an outcome variable (For example, changes in income inequality between two years) into factors that can be explained (such as the differences in partnering patterns or changes in female labour supply) and an unexplained residual component. The major limitation of this method is that it decomposes the mean values of the independent variables and the co-efficients from separate equations (in the example given, co-efficients from separate inequality equations for each year) and therefore only looks at a specific point along the distribution, the mean.[24] Finally, the DFL decomposition method allows for the construction of counterfactual densities using estimated weightings and therefore allows for reweighting of the original distribution density for those factors in question.

4.1. The Decomposition

We begin by defining the various observations with a vector that can be partitioned into three explanatory components. Let this vector be denoted as Ω consisting of equivalised household unit income (Y), a vector of observable and measurable characteristics (X) and the year (t). We can express the density of equivalised household unit income:

$$f_t(Y) = \int_{X \in \Omega_X} f(Y;X,t)dX \tag{4}$$

The density of equivalised household unit income in any given year t can be written as the integral of the conditional distribution of equivalised income over those measurable characteristics. Now applying the multiplicative properties of conditional distributions to the above equation, equation 4 therefore, can be represented by the following expression:

$$f_t(Y) = \int_{X \in \Omega_X} f(Y;X,t)\, f_X(X,t)dX \tag{5}$$

Equation 5 shows the equivalised household unit income distribution for each year as a function of a conditional and marginal probability density function. Equation 5 allows us to express any equivalised income distribution, conditional on the observable characteristics, in any given year. For example, if we were interested in the distribution of equivalised household unit income for 2003 with other observable characteristics measured in that same year, Equation 5 can be expressed as follows:

$$f_{03}(Y) = \int_{X \in \Omega_X} f(Y \mid X, t=03)\, f_X(X, t=03)dX \tag{6}$$

[24] For a more formal development of the Oaxaca Decomposition, please consult Greene (2003) pg54-55

Equation 6 represents the density of equivalised household unit income in 2003 with the distribution of measurable characteristics as they were in 2003. Effectively, this is an unadjusted distribution of household income in 2003 expressed in it conditional probability density form.

The major advantage of the above notation in equation 6 is that it allows the construction of counterfactual densities using observable characteristics from other years. Before beginning with a rudimentary example, let the vector of characteristics, X, assume one measurable attribute that represents mating patterns (A). Now let's say, for instance, we are interested in the density of equivalised income in 2003 had the assortative mating patterns from 1986 still prevailed. Now replacing the existing distribution with the mating attributes that would have prevailed in 1986, the desired counterfactual is obtained, with equation 7 reflecting this change.

$$f_{03}(Y) = \int_{A \in \Omega_A} f\left(Y \mid A, t = 03\right) f_A(A, t = 86) dA \qquad (7)$$

This expression represents the distribution that would be observed if the probability of mating patterns had retained its 1986 level with equivalised household unit income had the same distributional characteristics that were prevalent in 2003. DiNardo *et al* (1996) show that reweighting the original distribution can produce these types of counterfactual distributions. Rewriting equation 7

$$\begin{aligned} f_{03}(Y) &= \int_{A \in \Omega_A} f\left(Y \mid A, t = 03\right) \psi_A f_A(A, t = 03) dA \\ &= \int_{A \in \Omega_A} \psi_A f(Y, A \mid t = 03) dA \end{aligned} \qquad (8)$$

$$\text{where } \psi_A = \frac{f_A(A, t = 86)}{f_A(A, t = 03)} \qquad (9)$$

we can represent the counterfactual distribution as the original density of equivalised household unit income, with attributes as measured at 2003, reweighted by the function ψ_A. To obtain an estimate of the function ψ_A we use Bayes' theorem.[25]

[25] Recall Bayes' theorem with two discrete variables is given by the follow: $P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$

$$f_A(A \mid t=86) = \frac{f_A(A)P(t=86 \mid A)}{P(t=86)}$$

$$f_A(A \mid t=03) = \frac{f_A(A)P(t=03 \mid A)}{P(t=03)} \tag{10}$$

Substituting equation 10 into equation 9 and rearranging we get

$$\psi_A = \frac{P(t=86 \mid A)}{P(t=03 \mid A)} \frac{P(t=03)}{P(t=86)} \tag{11}$$

This function represents the change between 1986 and 2003.

4.2. Application of the Decomposition

A decomposition of the income distribution, we believe, is the best way to analyze our question of interest as it allows for us to specify specific groups that economic theory suggests affects the distribution of income. In doing so, it allows us to isolate the effects, especially for assortative mating, as it is our question of interest, and interpret their direct effect on income inequality. For the analysis we consider five groups in the decomposition of income inequality. In the following sequence we consider: (i) Labour force participation; (ii) Family characteristics; (iii) Mating patterns; (iv) Education levels and (v) Demographics.

It is important to pause here for a moment to discuss the ordering of the groupings mentioned above. Unlike simple OLS regression techniques, the order in which the groups enter the analysis matter. For example, group (i) is conditional on groups ii, iii, iv and v. In determining the order in which the groups enter the decomposition we use economic theory to define their place. We first consider labour force participation. Although labour force participation is dependant upon a number of factors, whether an individual chooses to work and the amount of time they choose to work depends strongly on a number of family characteristics. This is especially true for female labour supply, which also has further implications for income inequality. Similarly, family characteristics are a function of the choice to partner or not.

Dummy variables for the various education levels are included in the next grouping. Is it reasonable to conjecture that mating patterns or the decision to partner is conditional on one's education? Much of the sociology literature in the field suggests that educational institutions act like a quasi- singles market and it is the structure of these "marketplaces" that have a deterministic effect on who partners with whom. Even Becker (1973) suggests that educational facilities play an important role in the decision to marry. From an economic standpoint though, the returns to education include increased efficiency in the marketplace and the household. Since the returns to education can affect the productivity of market and household production, there is an influence on the decision to partner by the level of education an individual obtains. The former (marketplace productivity) affects the opportunity cost of marriage while increased household production increases the returns to marriage.

The final grouping introduces demographic variables including sex, age and born overseas, which are largely predetermined and are therefore considered the most exogenous of the five groupings. Therefore we start the decomposition with what we believe is the most endogenous of the five groupings, since labour force participation is conditional on the subsequent four groups that follow (For example, labour force participation is conditional on family characteristics, choice to mate, education levels and other demographics). We end with the most exogenous.

One of the other concerns of the analysis that has also been raised by other authors (Hyslop and Mare 2005; Daly and Valletta 2006) is that the results of the analysis might depend on the ordering of the variable groupings. The state an individual lives in might be dependant upon the work that individual does for example, or the size of one's family might depend upon what type of job the parents have or whether both parents participate in the labour market. The common approach to account for interdependence amongst the chosen groups is to reverse the order they enter the analysis. This method is adopted in Section 6 to test our results to sensitivity in the ordering of the groupings.

5. Results: Primary Decomposition

5.1. Changes between 1986 and 2003

We now turn our attention to the main part of our analysis, which is a primary order decomposition of the overall distribution of household unit incomes resulting from a set of factors, including assortative mating. We use a simple average counterfactual approach, analogous to the methodology used in Baron and Cobb-Clark (2006), to analyse the effect of the factor groupings on the distribution of incomes during the period 1986-2003.

The counterfactual decomposition is conducted using the five different groupings as mentioned in the previous section. These groups represent a vector of variables that include state of residence, job description and industry employed, number of dependant children, age and education to name a few.[26] In addition, interaction variables were also included in each group. Although not presented in Table 1, they are included to increase the precision of the probit model used to estimate the conditional counterfactual density weights.

The counterfactual density weights used in the analysis are presented in Table 6.[27] The counterfactual density weights are the conditional weights as obtained by a probit estimation, ψ_X, and are a product of the household unit sampling weights as provided by the ABS, θ_t. Row 1 of Table 6 shows the base distribution for 2003, which is the product of the unadjusted household income distribution in 2003 and the ABS household weights represented by θ_{03}. The subsequent rows represent the adjusted counterfactual 2003 distribution of equivalised household income. For example, row (l) identifies the base 2003 distribution as a product of the household weights, θ_{03}, and by the labour force participation ($\psi_{P|F,A,E,D}$), family characteristics ($\psi_{F|A,E,D}$) and mating patterns ($\psi_{A|E,D}$) as they were in 1986. The changes in the distribution can be shown in qualitative terms, using kernel density graphs, and in quantitative terms. Only the quantitative results are presented below.

[26] The other variables used in each group are presented in the summary statistics table (Table 1).

[27] For the sake of brevity only the counterfactual density weights have been presented in the table. In the analysis all counterfactual density weights are taken into account. For the case of 5 groups, there are possible 2^5 =32 counterfactual distributions. These counterfactuals including the original income distributions of 1986 and 2003 can be used to decompose inequality, using 5! = 120 different methods.

Table 6: Income meaures and Counterfactual Weights used in a 5 Factor Income Inequality Decomposition

	Counterfactual	Income Measure	Group 1 Lab. Partic	Group 2 Family Char	Group 3 AS Mating	Group 4 Education	Group 5 Demogr.	Weighting
	Base 2003 Distribution	Y	0	0	0	0	0	θ_{03}
(a)	03 Dist. with 1986 Labour Force Participation	Y_{03}	1	0	0	0	0	$\theta_{03}\psi_{P\|F,A,E,D}$
(b)	03 Dist. with 1986 Family Characterisitics	Y_{03}	0	1	0	0	0	$\theta_{03}\psi_{F\|A,E,D}$
(c)	03 Dist. with 1986 Mating Patterns	Y_{03}	0	0	1	0	0	$\theta_{03}\psi_{A\|E,D}$
(d)	03 Dist. with 1986 Education Levels	Y_{03}	0	0	0	1	0	$\theta_{03}\psi_{E\|D}$
(e)	03 Dist. with 1986 Demographics	Y_{03}	0	0	0	0	1	$\theta_{03}\psi_{D}$
(f)	(a) with 1986 Family Characterisitics	Y_{03}	1	1	0	0	0	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{F\|A,E,D}$
(g)	(a) with 1986 Mating Patterns	Y_{03}	1	0	1	0	0	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{A\|E,D}$
(h)	(a) with 1986 Education Levels	Y_{03}	1	0	0	1	0	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{E\|D}$
	(a) with 1986 Demographics	Y_{03}	1	0	0	0	1	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{D}$
(i)	(b) with 1986 Mating Patterns	Y_{03}	0	1	1	0	0	$\theta_{03}\psi_{F\|A,E,D}\psi_{A\|E,D}$
(j)	(b) with 1986 Education Levels	Y_{03}	0	1	0	1	0	$\theta_{03}\psi_{F\|A,E,D}\psi_{E\|D}$
	(b) with 1986 Demographics	Y_{03}	0	1	0	0	1	$\theta_{03}\psi_{P\|A,E,D}\psi_{D}$
(k)	(c) with 1986 Education Levels	Y_{03}	0	0	1	1	0	$\theta_{03}\psi_{A\|E,D}\psi_{E\|D}$
	(c) with 1986 Demographics	Y_{03}	0	0	1	0	1	$\theta_{03}\psi_{A\|E,D}\psi_{D}$
	(d) with 1986 Demographics	Y_{03}	0	0	0	1	1	$\theta_{03}\psi_{E\|D}\psi_{D}$
(l)	(f) with 1986 Mating Patterns	Y_{03}	1	1	1	0	0	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{F\|A,E,D}\psi_{A\|E,D}$
(m)	(f) with 1986 Education Levels	Y_{03}	1	1	0	1	0	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{F\|A,E,D}\psi_{E\|D}$
	(f) with 1986 Demographics	Y_{03}	1	1	0	0	1	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{F\|A,E,D}\psi_{D}$
(n)	(g) with 1986 Education Levels	Y_{03}	1	0	1	1	0	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{A\|E,D}\psi_{E\|D}$
	(g) with 1986 Demographics	Y_{03}	1	0	1	0	1	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{A\|E,D}\psi_{D}$
	(h) with 1986 Demographics	Y_{03}	1	0	0	1	1	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{E\|D}\psi_{D}$
(o)	(i) with 1986 Education Levels	Y_{03}	0	1	1	1	0	$\theta_{03}\psi_{F\|A,E,D}\psi_{A\|E,D}\psi_{E\|D}$
	(i) with 1986 Demographics	Y_{03}	0	1	1	0	1	$\theta_{03}\psi_{F\|A,E,D}\psi_{A\|E,D}\psi_{D}$
	(j) with 1986 Demographics	Y_{03}	0	1	0	1	1	$\theta_{03}\psi_{F\|A,E,D}\psi_{E\|D}\psi_{D}$
	(k) with 1986 Demographics	Y_{03}	0	0	1	1	1	$\theta_{03}\psi_{A\|E,D}\psi_{E\|D}\psi_{D}$
(p)	(l) with 1986 Education Levels	Y_{03}	1	1	1	1	0	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{F\|A,E,D}\psi_{A\|E,D}\psi_{E\|D}$
	(l) with 1986 Demographics	Y_{03}	1	1	1	0	1	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{F\|A,E,D}\psi_{A\|E,D}\psi_{D}$
	(m) with 1986 Demographics	Y_{03}	1	1	0	1	1	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{F\|A,E,D}\psi_{E\|D}\psi_{D}$
	(n) with 1986 Demographics	Y_{03}	1	0	1	1	1	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{A\|E,D}\psi_{E\|D}\psi_{D}$
	(o) with 1986 Demographics	Y_{03}	0	1	1	1	1	$\theta_{03}\psi_{F\|A,E,D}\psi_{A\|E,D}\psi_{E\|D}\psi_{D}$
	(p) with 1986 Demographics	Y_{03}	1	1	1	1	1	$\theta_{03}\psi_{P\|F,A,E,D}\psi_{F\|A,E,D}\psi_{A\|E,D}\psi_{E\|D}\psi_{D}$

Note: 1. Y refers to Real OECD equivalised gross income from salary and wages, with Y_{03} representing Real OECD equivalised income from salary and wages subject to changes in the 2003 Distribution 2. θ_t represents the sample weights as provided by the ABS.

Table 7 provides a quantitative representation of the primary decomposition. The rest of the paper uses this style of results for interpretation. Bootstrapped standard errors, based on 200 iterations, are presented along with the percentage contributions to the overall (raw) distributional effect. The standard errors are clustered at the household level. Significance of the statistics is indicated by an asteric. Table 7 presents the results of the decomposition of OECD equivalised household income, with assortative mating defined as a dichotomous variable.[28] Interpretation of the tables is as follows: A positive results indicates that the specific measured statistic has increased between 1986 and 2003. Conversely, a negative result indicates that the measured statistics has decreased over the two years. For example, row 1 of column 1 presents the raw gap for median household income. The result of 0.1571 is positive and significant and therefore indicates that the median level of real equivalised weekly household income has increased between 1986 and 2003.

Column 1 of Table 7 presents the raw gap difference between the distribution of 2003 and its counterfactual reweighted by characteristics from the 1986 income distribution. That is, it represents the overall difference between 1986 and 2003 for the presented income or income inequality measure. Inspection of Table 7 reveals that all of the dispersion measures presented have significantly increased during this time. The result of the 90-50 dispersion ratio (0.0503) indicates that the dispersion was greater in the upper half of the distribution than it was in the lower end, as indicated by the 50-10 dispersion ratio (0.0484). At the same time the results also indicate that the increase in dispersion in the middle section of the distribution (70-30 ratio) and at the lower end of the distribution (50-10 ratio) are quite similar.

The subsequent columns, columns 2 through to 6, present the contribution of the explanatory groupings to the overall raw gap. Any explanatory grouping can over or under explain the overall raw gap. Once again, bootstrapped standard errors are presented along with the overall percentage contribution to the total (raw gap) change.

---

[28] The primary decomposition uses the "old" OECD equivalence scale (sometimes known as the Oxford scale) where the first adult house member is assigned the value of 1 with all subsequent adults assigned a value of 0.7. Each child is assigned the value of 0.5. The robustness of the results is tested in Section 6 to differing equivalisation scales. For a further discussion of Equivalence Scales consult Lancaster and Ray (1998) and Blacklow and Ray (2000).

Table 7: Primary Decomposition of Changes in the Distribution of Real OECD Equivalised Household Unit Incon

	Measure	(1) Raw Gap	(2) LF Participation	(3) Family Characteristics	(4) Mating Patterns	(5) Education Levels	(6) Demographics	(7) Unexplained
	50th_Percentile	0.1571***	-0.0324***	-0.0024***	0.0007	0.0799***	-0.0020**	0.1134***
[1]	Std_Error	0.0103	0.0039	0.0007	0.0006	0.0043	0.0010	0.0090
	%	100%	-20.66%	-1.52%	0.44%	50.84%	-1.30%	72.20%
	50-10_Dif	0.0484***	-0.0020	0.0037**	0.0008	0.0422***	-0.0014	0.0050
[2]	Std_Error	0.0127	0.0039	0.0016	0.0007	0.0052	0.0010	0.0123
	%	100%	-4.12%	7.66%	1.67%	87.26%	-2.83%	10.36%
	70-30_Dif	0.0481***	0.0174***	0.0058***	0.0018***	0.0016	0.0015	0.0201*
[3]	Std_Error	0.0106	0.0038	0.0012	0.0007	0.0056	0.0011	0.0107
	%	100%	36.16%	12.03%	3.65%	3.24%	3.14%	41.78%
	90-50_Dif	0.0503***	0.017***	0.0025***	0.0034***	0.0050	-0.0009	0.0226***
[4]	Std_Error	0.0109	0.0036	0.0008	0.0010	0.0059	0.0010	0.0105
	%	100%	35.23%	5.02%	6.74%	9.94%	-1.77%	44.84%

** Primary Decomposition is estimated with the Assortative mating defined as a dummy variable indicating PAM or*

*Notes: 1. Standard Errors are constructed from bootstrapped statistics based on 200 repetions. 2. Percentages presented represent the percenta shares explained by the change in the various groups 3. Ratio dispersions are described as as the ratio between the relevant percentiles income. 4. *** indicates significance at 1% level, ** significance at 5%, * significan*

Column 7 presents the unexplained residual, that is, the difference between 1986 and 2003 that is let unexplained by our model. Columns 2 through 7, in summation, equal the raw gap (both statistic and percentage change) presented in column 1.

The first row presents the change in median household income and its decomposed components. The overall gap statistic is positive and indicates that there has been a real income increase in the median household income between 1986 and 2003. Given our question of interest, that is, how much of the increase in income inequality can be directly attributable to changes in assortative mating patterns, column 4 of Table 7 presents the effect assortative mating has had on income and income inequality. Both the contribution to the raw gap and its percentage effect are shown. The results indicate that between 1986 and 2003 changes in assortative mating have had a positive increase on the median household income. However, the magnitude of the effect, when compared to other groups in the decomposition, is relatively small. The statistic of 0.0007 only accounts for 0.44 percent of the total change. We therefore conclude that changing assortative mating patterns have had no substantial contribution to the increase in median equivalised household incomes.

Rows 2 – 4 present various dispersion ratios, which account for the lower (50-10), middle (70-30) and top end (90-50) of the distribution. The estimates listed in column 4 indicate the effect assortative mating has had on the key dispersion ratios presented. Changes in assortative mating explain a nominal amount, in terms of percentage, of the change in income inequality. The results show that around 1 – 6 percent of the increase in income inequality between 1986 and 2003, as measured by the 50-10, 70-30 and 90-50 ratios, are explained by changes in assortative mating patterns. The direction of the effect is in line with predictions of theoretical models in this area that suggest that PAM should increase inequality. Excluding the 50-10 ratio, the results for the 70-30 and 90-50 ratios account for 0.0018 and 0.0034 (respectively) of the raw gap. This represents 3.65 and 6.74 precent of the change in income inequality between 1986 and 2003 and indicates that there has been a “stringing out” of the middle and top end of the income distribution that is directly attributable to changing assortative mating patterns. The standards errors presented indicate that we can reject the null hypothesis of the co-efficient equalling zero. Assortative mating, therefore, has a significant effect on the changes in the middle and upper areas of the income distribution. Now turning our attention to the rest of the estimates presented in the remaining columns.

Column 2 of Table 7 lists the estimates of the effect of labour force participation on the raw gap. Overall, changing labour force participation patterns have had a significant and, compared to the other explanatory groupings listed, a substantial effect on the income and income inequality measures. Labour force participation can be attributed to a decreased in the median equivalised household income as indicated by the negative result of -0.0324. This accounts for a 21 percent understatement of the total effect. Excluding the 50-10 ratio, between one third and one half of the change in inequality, as measured by the ratio dispersions, can be accounted for by changes in labour force participation. Of the overall increase in the 90-50 ratio dispersion ratio (0.0503), 0.017 is due to changes in the labour force status between 1986 and 2003. This represents 33 percent of the overall change.[29] Similarly, labour force participation explains 36 percent of the increase in the 70-30 dispersion ratio. Although the results indicate that changing labour force participation has had similar effects on the upper and middle sections of the household income distribution, the results also indicate that its effect has been much smaller at the bottom end of the distribution (50-10 ratio) both in the magnitude of the statistic (-0.0020), its direction and as a percentage of the total effect (4 percent). Given the way we have defined labour force participation, we are unable to untangle the effects of increased female labour supply on the overall gap. Other authors (Daly and Valletta 2006) found that increased female participation decreased family level income dispersion in the United States. This would make for interesting further research to see if a similar pattern existed in an Australian context.

The third column presents the results for family characteristics. Family characteristics have decreased the median level of equivalised household income, explaining about 1 percent of the change in the median level income. At the same time changing family characteristics have had a substantial effect on income

[29] This result is in line with other work in the field on the distribution of Australian incomes. For example, Johnson and Wilkins (2004) also find that one third of the increase in inequality, as measured by the 90-50 ratio can be explained by changing labour force participation patterns.

inequality at the lower, middle and upper areas of the distribution only explaining 7.66 percent, 12.03 percent and 5.02 percent respectively of the overall change.

At this point it is worth pausing to look and the education grouping separately since it accounts for a significant amount of the overall changes in the various measures presented. Column 5 lists the effects of changing education patterns on income and income inequality. Changes in education levels account for 87 percent of the overall increase in income inequality as measured by the 50-10 ratio. Similarly, 50 percent of overall increase in the median household income is attributable to changing education levels. Interestingly though, those same patterns account for very little of the increase in income inequality at the middle and upper ranges of the distribution, which only account for 3.24 and 9 percent increase respectively. This indicates that increased educational attainment has had a much bigger impact than any other grouping in increasing inequality between those working aged individuals at the very bottom of the income distribution.

Column 6 presents the results of the demographic grouping. There is very little of interest to explain within this group. Compared to the other explanatory groupings, the impact of changing demographics on income and income inequality has been neutral. The results indicate that demographics have had little significant impact on the median income level with demographics understating median level income by 1 percent. A similar story is listed with the ratio dispersion. Demographics only explain 2.83 percent and 3.14 percent of the increase in the 50-10 and 70-30 dispersion ratios respectively.

Finally, with 72 percent of the increase in the median income level unexplained and between 10 percent and 44 percent of the change in income inequality also unaccounted for, the contribution of the unexplained or residual factors, as presented by the final column, suggests that our choice of explanatory groups explains between 65 percent and 90 percent of the changing inequality between 1986 and 2003. Although the unexplained residuals suggest that a large part of the change in income and income inequality between the two years is still left unaccounted for, especially at the top end, we believe that our model of choice is still effective in explaining many factors that contribute to the overall effect. As mentioned previously, the result statistics and percentage changes in columns 2 through 7 add up to the raw gap statistic in column 1. If the individual grouping result has a different sign from the overall gap then the percentage change is presented as a negative change (for example, see Family Characteristics). The unexplained residual can therefore be overstated by the fact that some group characteristics move in the opposite direction to the raw gap. In evaluating the overall effectiveness of the model is best to look at the absolute value of the percentage explained. With this in mind, we see that assortative mating patterns across the different income and income inequality measures explain a moderate amount of the change between 1986 and 2003. Conversely, labour force participation and changing education levels, in absolute terms explain 20-36 percent and 50-87 percent of the raw gap respectively.

6. Robustness Checks

As an alternative analysis to the primary decomposition, we investigate the sensitivity of the results to differing equivalisation methods and definitions of assortative mating. The results are presented below in Table 8 and Table 9

6.1. Differing Equivalising methods

Table 8 presents the results to differing methods of equivalising income. The three different methods are: 1) The "Modified" OECD Scale, 2) The Square Root Scale and 3) Per Capita scale.[30] For the sake of brevity we only report the results on the effect of assortative mating, therefore, they are directly comparable to column 4 of Table 7. Column 1 of Table 8 presents the raw gap of the decomposition with "modified" OECD equivalised income with Column 2 showing the effect of assortative mating on that same raw gap. Again all income and income equality measures show a significant increase between 1986 and 2003. The results in column 2 indicate that changes in assortative mating patterns have increased income inequality across all measures, significantly in the middle and top ends of the distribution. The 50-10 ratio indicates that assortative mating increased inequality in the bottom end of the distribution

[30] The "modified" OECD scale assigns a value of 1 to the first adult household member with each subsequent adult assigned the value of 0.5. Each child is assigned the value of 0.3. The square root method is a scale that divides household income by the square root of the household size. Per capita scale divides the household income by the size of the household and therefore assumes no economies of scale to household size.

although not significantly, with it explaining 2 percent of the total increase in income inequality between 1986 and 2003. In terms of magnitude, assortative mating has had the greatest impact at the top end of the distribution with the 90-50 ratio indicating that assortative mating accounts for 6.5 percent of the total increase in the top end of the distribution.

Column 3 and 4 of Table 8 present the results of the decomposition of “Square Root” equivalised income and assortative mating’s impact on the overall change. The results are quite interesting and in contrast to the previous equivalisation results. Although the results indicate that for each dispersion measure, income inequality increased, for the 50-10 and 70-30 dispersion ratios, assortative mating explains nothing of the increase in the income inequality in terms of statistical significance. We therefore conclude that assortative mating has no impact on the lower end and middle areas of the income distribution. However further inspection reveals that at the top end of the distribution there has been a significant increase in income inequality directly attributable to this definition of assortative mating. The 90-50 ratio indicates that assortative mating significantly accounts for 1.96 percent of the increase in income inequality in the top half of the distribution.

Table 8: Robustness Checks - Summary of Assortative Mating Results given changes in equivalence scale

	Statistic	(1) Raw Gap	(2) AS Mating#	(3) Raw Gap	(4) AS Mating##	(5) Raw Gap	(6) AS Mating###
	Median	0.1597***	0.0008	0.1204***	-0.0017***	0.1504***	0.0002
[1]	CI	0.0091	0.0006	0.0082	0.0007	0.0098	0.0006
	%	100%	0.50%	100%	-1.40%	100%	0.16%
	50-10_Dif	0.0458***	0.0009	0.0138	0.0010	0.0588***	0.0003
[2]	CI	0.0137	0.0009	0.0135	0.0011	0.0164	0.0009
	%	100%	2.05%	100%	7.10%	100%	0.43%
	70-30_Dif	0.0626***	0.0015**	0.0480***	0.0006	0.0448***	0.0008
[3]	CI	0.0107	0.0007	0.0114	0.0007	0.0099	0.0008
	%	100%	2.47%	100%	1.17%	100%	1.88%
	90-50_Dif	0.0579***	0.0038***	0.0631***	0.0012***	0.0366***	0.0016**
[4]	CI	0.0123	0.0011	0.0122	0.0008	0.0120	0.0008
	%	100%	6.54%	100%	1.96%	100%	4.46%

*Notes: 1. Standard Errors are constructed from bootstrapped statistics based on 200 repetitions. 2. Percentages presented represent the percentage shares explained by the change in the various groups 3. Ratio dispersions are described as as the ratio between the relevant percentiles of real equivalised income. 4. *** indicates significance at 1% level, ** significance at 5%, * significance at 10%*

 # Modified OECD Equivalised Income

Square Root Equivalised Income

Per Capita Equivalised Income

Columns 5 and 6 present the results under a per capita equivalising method. Although not very common in the literature, since it assumes no economies of scale to household size, it at least provides us with an upper bound limit on the effect of assortative mating at the extreme end of household equivalence elasticities.[31] Excluding the 90-50 ratio, Column 6 indicates that although assortative mating has increased the median income and inequality in the lower and middle parts of the income distribution, its overall effect is not significant. In addition, in terms of its percentage explanation of the raw gap (column 5), assortative mating accounts for very little of the raw gap, ranging between 0.43 and 1.88 percent. However, the 90-50 ratio indicates that assortative mating has a significant effect on the increase in income inequality at the upper end of the income distribution with assortative mating accounting for 4.46 percent of the overall increase in income inequality. Although slightly smaller in magnitude, this reinforces our results under the “old” OECD, “modified” OECD and “Square Root” equivalisation methods alike.

[31] Effectively equivalence scales can be expressed as an “equivalence elasticity”. That is, the size at which the economic production and consumption of the household change with its size. The elasticity can range from 0 for unadjusted income through to 1 in the extreme case of per capita equivalisation.

6.2. Changing definition of Assortative mating

We also investigate whether the patterns that have been identified were also prevalent when we controlled for NAM and single households. Recall from Section 3 that our assortative mating group only accounts for PAM individuals or otherwise, since the economic theory in this area predicts that it is PAM that is more likely to widen the distribution of a measurable trait than other forms of mating.

Table 9: Robustness Checks - Summary of Assortative Mating Results given changes the definition of assortative mating

	Statistic	(1) Raw Gap	(2) AS Mating#	(3) Raw Gap	(4) AS Mating ##
	Median	0.1571***	-0.0002	0.1571***	0.0029**
[1]	CI	0.0103	0.0013	0.0103	0.0014
	%	100%	-0.12%	100%	1.87%
	50-10_Dif	0.0484***	0.0002	0.0484***	0.0019*
[2]	CI	0.0127	0.0008	0.0127	0.0010
	%	100%	0.43%	100%	3.85%
	70-30_Dif	0.0481***	0.0021***	0.0481***	-0.0003
[3]	CI	0.0106	0.0008	0.0106	0.0007
	%	100%	4.30%	100%	-0.61%
	90-50_Dif	0.0503***	0.0035**	0.0503***	-0.0009
[4]	CI	0.0109	0.0013	0.0109	0.0007
	%	100%	6.93%	100%	-1.69%

*Notes: 1. Standard Errors are constructed from bootstrapped statistics based on 200 repetitions. 2. Percentages presented represent the percentage shares explained by the change in the various groups 3. Ratio dispersions are described as as the ratio between the relevant percentiles of real equivalised income. 4. *** indicates significance at 1% level, ** significance at 5%, * significance at 10%*

 # Assortative Mating controls for those partnerships which are PAM, NAM and Singles

Assortative Mating is defined as the product of an individuals distances from a gender specific median to their spouses distance from a gender specific median

Results are presented in Columns 1 and 2 of Table 9 with column 1 presenting the raw gap and column 2 indicating the assortative mating groups influence on the raw gap. The results are relatively insensitive to this change in specification. When we control for all those individuals who are NAM and not partnered we find that mating patterns now explain 6.9 percent of the change in the top end ratio (90-50). This result is slightly larger than tat found in the primary decomposition, which indicated that assortative mating patterns explained 6.73 percent of the increase in the 90-50 ratio. In addition, although the median level income increases over the two years, there is no significant effect of changing mating patterns on its increase. For the lower end and middle section of the income distribution, assortative matings effect on the increase in income inequality is less than under the primary decomposition, although the assortative mating group does indicate that it has had a positive effect on income inequality.

Finally, we investigate the sensitivity of the primary decomposition results presented in Section 5.1 to one last change in the definition of assortative mating. As an alternative method of determining assortative mating we create a new variable, which is defined as the product of the distance of each partners education level from the gender specific median at time t. A formal presentation of this is given in Section 3. The results are presented in columns 3 and 4 of Table 9. The results indicate that the findings are sensitive to this change in the definition of assortative mating with assortative mating having no significant effect on income inequality in the middle and upper ends of the income distribution.

Not only is the magnitude of the effect, as a percentage of the raw gap, smaller (6.78 percent in our primary decomposition compared to 0.61 percent for the 90-50 ratio) but also the direction of the effect is now in the opposite direction. Although Table 1 reveals that, on average, individuals display PAM (by this definition) on education across 1986 and 2003, the magnitude of the change in assortative mating, however, is small. Since the standard errors indicate that the co-efficients are not significant we therefore conclude that this definition of assortative mating has had no significant impact on the increase in income inequality as measured by the 90-50 and 70-30 ratios from 1986 to 2003.

Interestingly though, for the 50-10 ratio, we find the largest significant effect on income inequality under this definition of assortative mating. Assortative mating accounts for 3.85 percent of the increase in income inequality at the lower end of the distribution.

6.3. Reverse Order Decomposition

As discussed in section 3 of the paper, the contributions of the explanatory groups of choice depend on the order in which they enter the decomposition analysis. However, there could be good justification in assuming there is some sort of joint determination or joint causality amongst the groups. For example, the amount of time an individual dedicates to the labour force may influence mating patterns, especially given that, on average, Australian workers are working more hours than in the past.
Or similarly, the size of one's family depends on what type of job the parents have or whether both parents participate in the labour market. The common approach to account for interdependence amongst the chosen groups is to reverse the order they enter the analysis. We therefore conduct this reverse order decomposition for the period 1986 -2003, where we consider: Demographics, Education, Mating patterns, Family Characteristics and Labour Force Participation, in this order. Results are presented in Table 10.

The impact of changing mating patterns has increased in magnitude, especially with the 90-50 inequality ratio and provides us with an upper boundary of assortative matings effect on income inequality. Compared to the primary decomposition, the effects of PAM on inequality, as measured by the 90-50 ratio, have increased from 6.74 percent to 10.56 percent. In addition the percentage effect of assortative mating on the raw gap for the 70-30 ratio has double from 3.65 percent to 7.25 percent. Overall, the magnitude of assortative matings effect on all income inequality measures is larger. Therefore, the reverse order decomposition indicates that assortative mating has an increasing effect on the measures presented, which is consistent with the primary decomposition.

The increased effect of the assortative mating group in the reverse order decomposition, however, is at the expense of the other groups in the decomposition. The impact of changing labour force participation, when taken last the ordering, does not have as large an impact on the raw gap. In addition, its effect on the raw gap has reversed and now understates the raw gap figure. Compare this to when it is first in the ordering, where labour force participation positively explained around one third of the increase in the dispersion measures between 1986 and 2003. Demographics,

Table 10: Reverse Order Decomposition of Changes in the Distribution of OECD Equivalised Household Unit Income (1986-2003)#

	Measure	(1) Raw Gap	(2) Demographics	(3) Education Levels	(4) Mating Patterns	(5) Family Characterisitics	(6) LF Participation	(7) Unexplained
	50th_Percentile	0.1571***	-0.0054***	0.0495***	0.0005	-0.0051***	0.0042	0.1134***
[1]	Std_Error	0.0103	0.0014	0.0036	0.0008	0.0010	0.0041	0.0090
	%	100%	-3.43%	31.49%	0.33%	-3.23%	2.65%	72.20%
	50-10_Dif	0.0484***	-0.0031**	0.0441***	0.0018*	0.0018	-0.0012	0.0050
[2]	Std_Error	0.0127	0.0013	0.0043	0.0010	0.0019	0.0023	0.0123
	%	100%	-6.37%	91.13%	3.64%	3.82%	-2.58%	10.36%
	70-30_Dif	0.0481***	0.0022	0.0238***	0.0035***	0.0075***	-0.009***	0.0201*
[3]	Std_Error	0.0106	0.0015	0.0050	0.0010	0.0014	0.0025	0.0107
	%	100%	4.55%	49.54%	7.25%	15.60%	-18.71%	41.78%
	90-50_Dif	0.0503***	0.0006	0.0229***	0.0053***	0.0039***	-0.005*	0.0226**
[4]	Std_Error	0.0109	0.0015	0.0055	0.0014	0.0010	0.0026	0.0105
	%	100%	1.16%	45.56%	10.56%	7.80%	-9.92%	44.84%

 # Reverse Order Decomposition is estimated with the Assortative mating defined as a dummy variable indicating PAM or otherwise

*Notes: 1. Standard Errors are constructed from bootstrapped statistics based on 200 repetions. 2. Percentages presented represent the percentage shares explained by the change in the various groups 3. Ratio dispersions are described as as the ratio between the relevant percentiles of Real OECD equivalised income. 4. *** indicates significance at 1% level, ** significance at 5%, * significance at 10%*

although now significant on two of the dispersion measures, still explain very little of the increase in income inequality. Education levels are still quite important with their percentage magnitude decreasing slightly from the primary order decomposition. Overall, the unexplained factors remain consistent between both the primary and reverse order decompositions.

7. Conclusion

Over the past two decades there has been increased concern about rising income inequality here in Australia. Many authors have charted its course and all agree income inequality has increased. There have been many reasons given to why income inequality has increased in the last 20 years, with this work confirming some of the results found in previous work. However, the work also makes one important departure. Implications from a number of theoretical constructs show clearly increases in assortative mating among traits, including wages and income, can have important implications for inequality. As noted from the outset, the prevalence of marriages that exhibit assortative mating can increase economic inequalities between household units. It is this finding that we are most concerned in empirically testing.

We focus on the period of 1986-2003, which entails a period of large inequality growth in Australia (1980's) followed by a relatively flat period of increasing inequality (1990's). Using a semi parametric weighted density estimation framework, our results indicate that changing patterns of assortative mating over the same period account for very little of the increase in income inequality in Australia. Our best estimates reveal that the counterfactual impact of assortative mating accounts for 1 – 6 percent of the increase in inequality between the years 1986 and 2003, with much of the movement being in the top end of the income distribution.

Estimates were also undertaken using differing methods of income equivalisation. Although the impact and significance of some of the findings changed, in general, the findings confirm our results under the primary decomposition with 2-6 percent of income inequality movements accounted for by changing assortative mating patterns, once again with much of the movement occurring in the top end of the income distribution.

The results were mixed to the definition of assortative mating used. When expanding our definition of assortative mating to capture those individuals who were NAM and single, our assortative mating group accounted for between 4-7 percent of the increase in income inequality over the 17 years to 2003, confirming our results from the primary decomposition and different equivalising methods. In contrast, however, our second definition of assortative mating indicated that assortative mating had no influence on household income inequality.

In using economic theory, we determine four other explanatory grouping to help explain the increase in inequality over the past 17 years. We also find, in line with results from other authors, that changing patterns in labour force participation explain roughly one third of the increase in income inequality. It is difficult to say whether the changes in labour force participation have been driven largely by increases in female labour supply, as some authors suggest, but this would make for interesting further research. In addition, changing education levels explain a large majority of the increase in income inequality over the two years with it also accounting for one half of the increase in the median level of household income. The results also indicate that the effect education has had on the different ends of the income distribution is quite different. Education accounts for a much of the increase in income inequality at the lower end of the income distribution, with only a comparatively nominal increase in inequality noticed in the middle and top end of the income distribution. Changing family characteristics have a significant effect in income and income inequality, which is in line with findings in the United States. Finally, demographics have little explanatory power, with the result robust to its position in the regression equation.

References

Baron, J and Cobb-Clark, D (2006): “A Distributional Analysis of Gender Wage Differences in Australia”, Unpublished Working Paper

Barrett, G. Crossley, T and Worswick, C (2000): “ Consumption and Income Inequality in Australia”, *The Economic Record* 76: 116-138

Blacklow, P and Ray, R (2000): “A Comparison of Income and Expenditure Inequality Estimates: The Australian Evidence, 1975-1976 to 1993-1994”, *The Australian Economic Review* 33: 317-329

Becker, G S (1973): “A Theory of Marriage: Part 1”, *The Journal of Political Economy* 81: 813-846

Becker, G S (1991): A Treatise on the Family, Cambridge: Harvard University Press

Borland, J and Wilkins, R (1996): “Earnings Inequality in Australia”, *The Economic Record* 72: 7-23

Burdett, K and Coles, M (1997): “Marriage and Class”, *The Quarterly Journal of Economics* 12: 141-168

Butcher, K and DiNardo, J (2002): “The Immigrant and native-Born Wage Distributions: Evidence from United States Censuses”, *Industrial and Labor Relations Review* 56: 97-121

Çelikaksoy, A. Nielsen, H S and Verner, M (2003): “Marriage Migration: Just another case of positive assortative matching?”, *Unpublished working paper*, Aarhus School of Business

Conceição, P and Ferreira, P (2000): “The Young Person’s Guide to the Theil Index: Suggesting Intuitive Interpretations and Exploring Analytical Applications”, *UTIP Working Paper Number 14*, The University of Texas at Austin

Daly, M and Valletta, R (2006): “Inequality and poverty in the United States: The Effects of rising Dispersion of men’s Earnings and Changing Family Behaviour, *Economica* 73: 75-98

DiNardo, J. Fortin, N and Lemieux, T (1996): “Labor Market Institutions and the Distributions of Wage, 1973-1992: A Semi parametric Approach”, *Econometrica* 64: 1001-1044

Elifson, K (1982): Fundamentals of Social Statistics, Addison-Wesley Publishing Company

Fernandez, R. Guner, N and Knowles, J (2001): “Love and Money; A Theoretical and Empirical Analysis of Household Sorting and Inequality”, *NBER Working Paper Series* 8580

Fernandez, R and Rogerson, R (2001): “Sorting and Long Run Inequality”, *The Quarterly Journal of Economics* 116: 1305-1341

Greene, W H (2003): Econometric Analysis 5th Edition, New Jersey: Prentice Hall

Halpin, B and Chan, T W (2002): “Union Dissolution in the United Kingdom”, *International Journal of Sociology* 32: 76 -93

Harding, A (1997): “The Suffering Middle: Trends in Income Inequality in Australia, 1982-1993/4”, *The Australian Economic Review* 30: 341-58

Harding, A and Greenwell, H (2002): “Trends in Income and Expenditure Inequality in the 1980’s and 1990’s: A Re-Examination and further results” *NATSEM Discussion Paper Series 2002 No.57*

Hays, W L (1988): Statistics, New York: The Dysden Press Saunders College Publising

Hyslop, D and Mare, D (2005): “Understanding New Zealand’s Changing Income Distribution, 1993-1998: A Semi parametric Analysis”, *Economica* 72: 469-195

Johnson, D and Wilkins, R (2004): “Effects of changes in family Composition and Employment Patterns on the Distribution of Income in Australia: 1981-1982 to 1997-1998”, *The Economic Record* 80: 219-238

Kremer, M (1997): “How much does Sorting Income Inequality?”, *The Quarterly Journal of Economics* 112: 115-139

Lam, D (1988): “ Marriage Markets and Assortative Mating with Household Public Goods: Theoretical results and Empirical Implications”, *The Journal of Human Resources* 23: 462-487

Lancaster, G and Ray, R (1998): “Comparison of Alternative Models of Household Equivalence Scales: The Australian Evidence on Unit Record Data”, *The Economic Record* 74: 1-14

Leigh, A (2005): “Deriving Long-Run Inequality Series from Tax Data”, *The Economic Record* 81: 58-70

Marks, G.N, Headey, B and Wooden, M (2005): “Household Wealth in Australia: It’s Components, Distribution and Correlates”, *Journal of Sociology* 41: 47-68

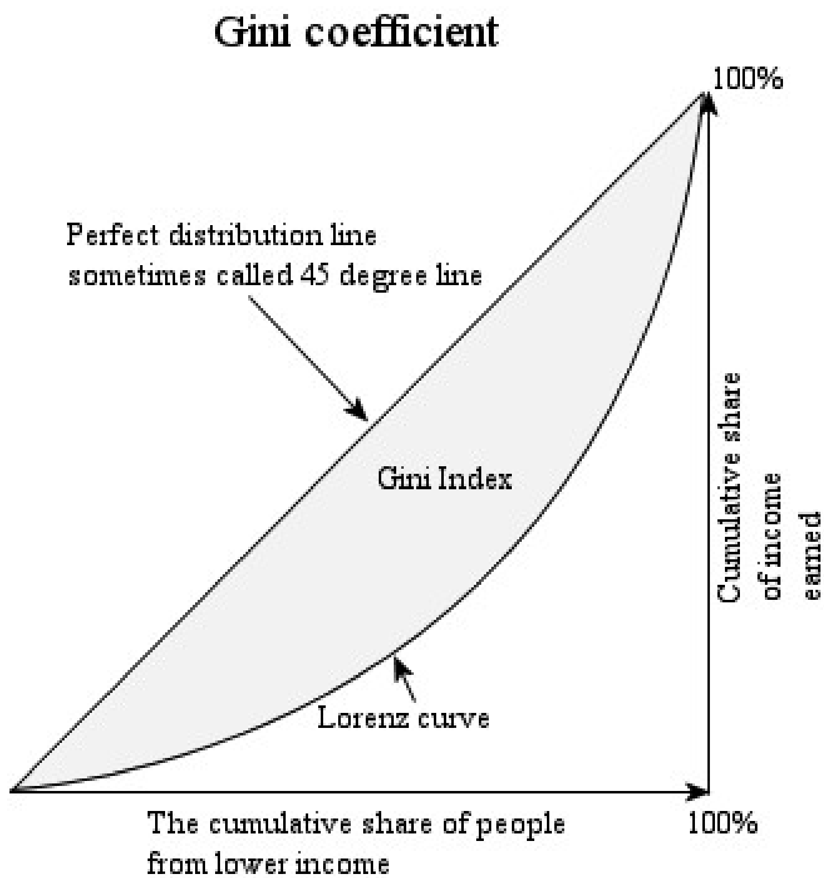
Oaxaca, R (1973): "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review* 14: 693-709.

STATA Reference Manual Release 7 Volume 2 H-P, College Station Texas, STATA Press

Tsai, SL and Chung-Ming, K (2004): “A Quantile Regression Analysis on Family Income Determination with Mating Effects”, Unpublished Working paper

Appendix

- *Graphical Representation of the GINI co-efficient*



Graphical representation of Income Equality and Income measures, Per capita equivalised