Using Relative Distribution Software

Michele L. SHAFFER and Mark S. HANDCOCK

Relative distribution methods are a nonparametric statistical approach to the comparison of distribution. These methods combine the graphical tools of exploratory data analysis with statistical summaries, decomposition, and inference.

This technical report describes how relative distribution methods can be implemented using the Relative Distribution Methods Software, a freeware collection of Splus functions. The software is illustrated by implementing procedures described in the paper "Relative Distribution Methods" by Mark S. Handcock and Martina Morris, *Sociological Methodology*, Vol 28, July 1998.

1. INTRODUCTION

In social science research, differences among groups or changes over time are a common focus of study. While means and variances are typically the basis for statistical methods used in this research, the underlying social theory often implies properties of distributions that are not well captured by these summary measures. Consider some of the current controversies regarding growing inequality in earnings, racial differences in test scores, socio-economic correlates of birth outcomes, and the impact of smoking on survival and health. The distributional differences that animate the debates in these fields are complex. They comprise the usual mean-shifts and changes in variance, but also more subtle comparisons of changes in the upper and lower tails of the distributions. Survey and census data on such attributes contain a wealth of distributional information, but traditional methods of data analysis leave much of this information untapped.

Handcock and Morris (1998) (HM) present methods for full comparative distributional analysis. The methods are based on the relative distribution, a non–parametric complete summary of the information required for scale–invariant comparisons between two distributions. The relative distribution provides a general integrated framework for analysis: a graphical component that simplifies exploratory data analysis and display, a statistically valid basis for the development of hypothesis–driven summary measures, and the potential for decomposition that enables one to examine complex hypotheses regarding the origins of distributional changes within and between groups.

Mark S. Handcock is Professor of Statistics and Sociology, Department of Statistics, University of Washington, Seattle, WA 98195-4320 (E-mail: handcock@stat.washington.edu) and; Michele L. Shaffer is a doctoral student, Department of Statistics, Pennsylvania State University, University Park, PA 16802 (E-mail: shaffer@stat.psu.edu) and; The authors wish to thank Annette D. Bernhardt, Jeffrey S. Simonoff and Martina Morris for comments that have greatly improved this paper.

The data used throughout are from two cohorts of the National Longitudinal Survey (NLS), one initiated in 1966 and the other in 1979. These cohorts are referred to as the original and recent cohorts, respectively. The distributions of wage growth in the two cohorts are examined. Specifically, the growth profile of "permanent wages" is analyzed to study the question of wage mobility. A development of the estimation of these permanent wages and their relevance to the study of wage mobility is given in HM. For the purposes of this report, we can regard the permanent wages as measurements on two groups that we wish to compare.

The software described here can be obtained at `http://lib.stat.cmu.edu/S/reldist`, or by sending the email message "send reldist from S" to `statlib@lib.stat.cmu.edu`. In addition, information is available from the second author's home page (`http://www.stat.washington.edu/handcock/RelDist`). Additional technical reports of interest can also be found here.

In the following sections we will reconstruct the analyses given in HM without discussing their substantive interpretation. To fully understand the use of the software it would be useful to have a copy of HM available for reference. The appropriate function calls and associated code are given for the methods described by section. In Section 2, the standard approach to comparing the two distributions is presented. In Section 3, the relative CDF and PDF of permanent wage growth in the original and recent NLS cohorts are constructed. In Section 4, the relative distribution of permanent wage growth in the two cohorts is decomposed into the impact of changes in medians and changes in shape. In Section 5, summary statistics for the location/shape decomposition of the relative distribution of wage gains are computed. In Section 6, an example of covariate adjustment is provided, adjusting the relative distribution of permanent wage growth for changes in educational composition between the two cohorts. And finally in Section 7, the code necessary for reproducing a discrete level contrast example and an additive decomposition example is given.

2. DENSITY ESTIMATION

The standard approach to comparing the original and recent NLS cohort distributions involves looking at the summary statistics and plotting the probability density functions (PDFs) and Lorenz curves. The built in `S` function used for estimating the PDFs is `density`. The plots of the PDFs given in Figure 1 (a) are generated by

```
kwidth <- 0.2597
dens1 <- density(schpermwage1, n = 500, width=kwidth)
plot(x = (dens1$x), y = dens1$y, type = "l",
  xlab = "change in log permanent wage", ylab = "density", axes = F,
```

```
  xlim = c(-1, 3), ylim=c(0,1.2))
title(main="(a)",cex=0.6)
axis(side = 1)
axis(side = 2)
fig1legend <- list(x=c(1.2,1.2),y=c(1.2,1.2))
legend(fig1legend,lty=1:2,cex=0.5, bty="n",
  legend=c("Original cohort","Recent cohort"))
dens2 <- density(schpermwage2, n = 500, width=kwidth)
lines(x = (dens2$x), y = dens2$y, type = "l",lty=2)
```

From the plot of the PDFs, we see that the recent cohort experienced smaller average wage gains, these gains were more variable, and the frequency of low wage gains was much greater for the recent cohort.

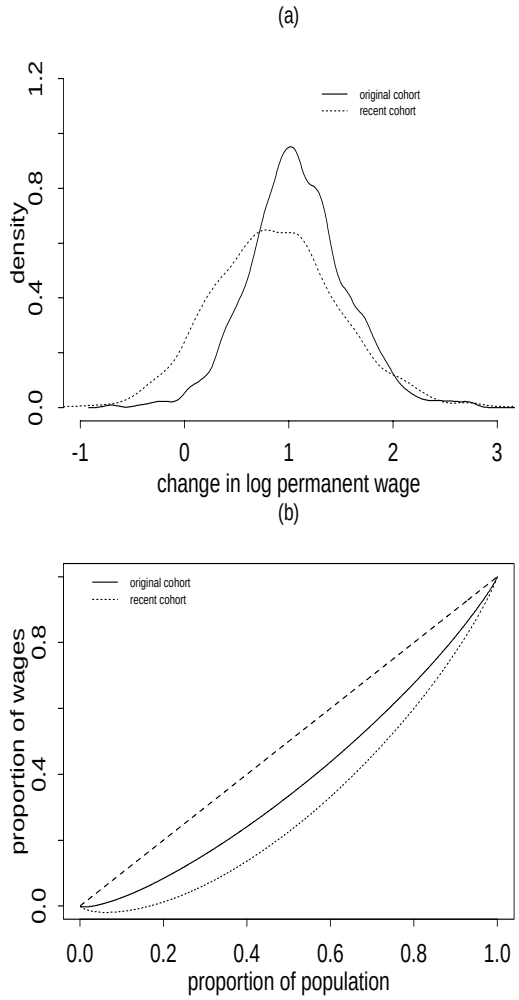

Figure 1. The distributions of permanent wage growth in the original and recent NLS cohorts. (a) PDF overlays for each cohort; (b) Lorenz curves for the PDFs.

Lorenz curves are a standard method used for inequality comparison. The plots of the Lorenz curves given in Figure 1 (b) are generated by

```
swage1 <- sort(recent$chpermwage)
swage2 <- sort(original$chpermwage)
xout <- (0:1000)/1000
alpha <- seq(along=swage2)/length(swage2)
galpha <- cumsum(swage2)/sum(swage2)
fn1 <- approx(x=alpha,y=galpha,xout=xout)
plot(x = alpha, y = galpha, type = "l",
  xlab = "proportion of population", ylab = "proportion of wages",
  ylim=c(0,1.0))
legend(x=c(0,0),y=c(1.03,1.03),lty=1:2,cex=0.5, bty="n",
  legend=c("original cohort","recent cohort"))
```

Here we see that the Lorenz curve for the recent cohort lies uniformly below that of the original cohort which indicates that there is more inequality in the distribution of recent wage gains.

3. THE RELATIVE DISTRIBUTION

To do a full distributional comparison based on the relative distribution, we look at the PDF and CDF of the relative distribution. Both preserve all of the information necessary to compare the two distributions. If the two distributions are identical, then the CDF of the relative distribution is a 45° line and the PDF of the relative distribution is a uniform PDF.

To obtain the relative CDF in Figure 2 (a) we use

```
fig1 <- reldist(y=recent$chpermwage,yo=original$chpermwage,
        yowgt=original$wgt,ywgt=recent$wgt,
        cdfplot=T,
        smooth=0.4,
        yolabs=seq(-1,3,by=0.5),
        ylabs=seq(-1,3,by=0.5),
        cex=0.8,
```

```
        ylab="proportion of the recent cohort",
        xlab="proportion of the original cohort")
title(main="(a)",cex=0.6)
```

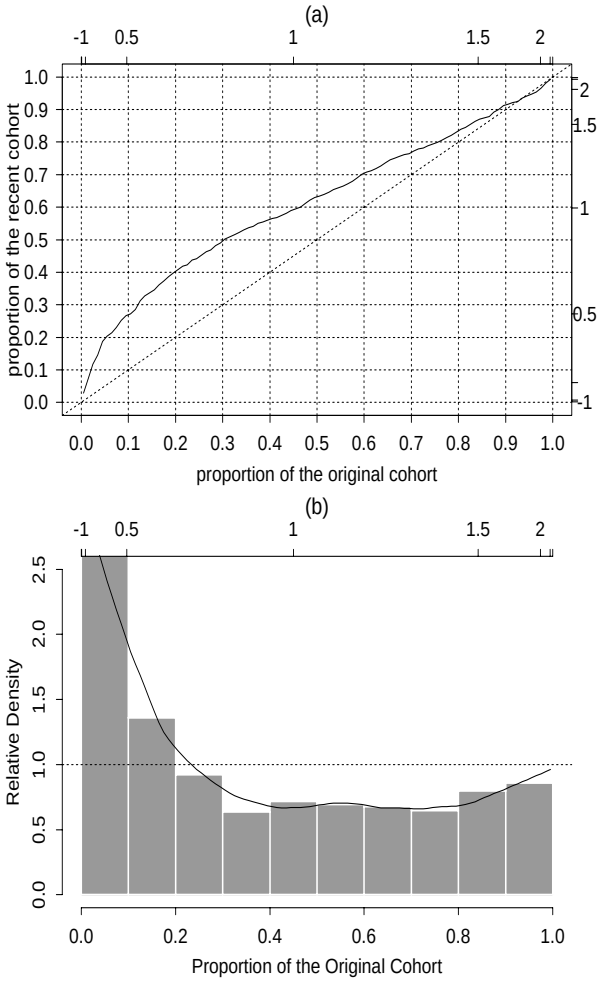

Figure 2. The relative distribution of permanent wage growth in the original and recent NLS cohorts: (a) the relative CDF; (b) the relative PDF. A decile bar chart is superimposed on the density estimate. The upper and right axes are labeled in permanent differences in log wages.

The option `cdfplot = T` is used to obtain a plot of the CDF rather than the (default) density. The options `yowgt` and `ywgt` are used to assign a vector of weights to the reference distribution and comparison distribution, respectively. The `smooth` option identifies the degree of smoothness required in the fit. Specifying higher values of `smooth` leads to smoother curves, while specifying lower values leads to closer fits to the observed data. The relative PDF in Figure 2 (b) is produced with

```
fig1 <- reldist(y=recent$chpermwage,yo=original$chpermwage,
        yowgt=original$wgt,ywgt=recent$wgt,
        bar=T,
        smooth=0.4,
```

```
            yolabs=seq(-1,3,by=0.5),
            ylim=c(0,2.5),cex=0.8,
            ylab="Relative Density",
            xlab="Proportion of the Original Cohort")
    title(main="(b)",cex=0.6)
```

Here the option `bar=T` is used to superimpose a barplot on the relative density estimate. Figures 3 (a) and (b) show the effects of increasing the `smooth` option used in Figures 2 (a) and (b) by specifying `smooth = 1.2`.

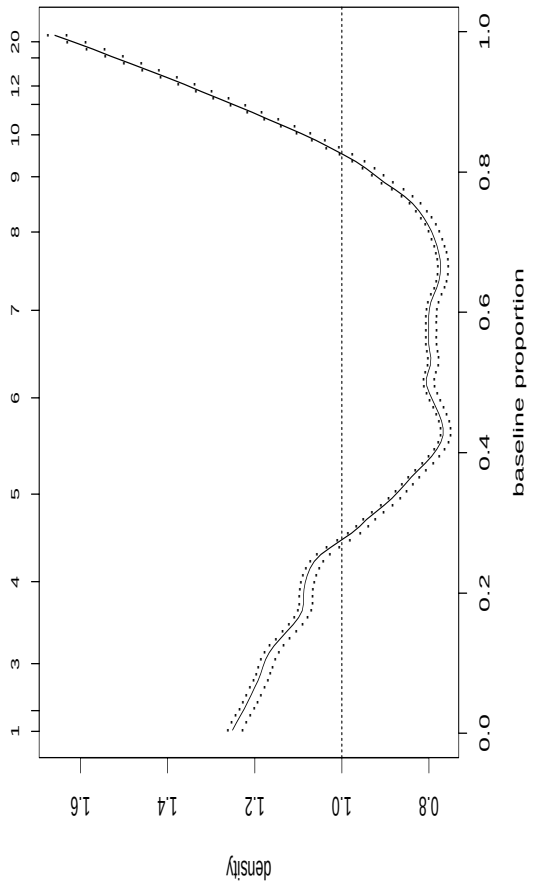

Figure 3. The relative distribution of permanent wage growth in the original and recent NLS cohorts: (a) the relative CDF; (b) the relative PDF. A decile bar chart is superimposed on the density estimate. The upper and right axes are labeled in permanent differences in log wages.

4. DECOMPOSING THE RELATIVE DISTRIBUTION

In this section we decompose the overall relative distribution into two component relative distributions which depict differences in location and shape. Figure 4 displays the median and shape decomposition of the relative distribution of weight gains and is generated by

```
par(err=-1)

par(mfrow=c(1,3))

g10 <- reldist(y=recent$chpermwage, yo=original$chpermwage,

       ywgt=recent$wgt, yowgt=original$wgt,

       smooth=0.4,

       yolabs=seq(-1,3,by=0.5),

       ylim=c(0.5,3.0),

       bar=T,

       xlab="proportion of the original cohort")

title(main=paste("(a) entropy = ",format(g10$entropy,digits=3)),cex=0.6)

abline(h=1,lty=2)

g1A <- reldist(y=recent$chpermwage, yo=original$chpermwage,

       ywgt=recent$wgt, yowgt=original$wgt,

       matchshape=T,

       bar=T,

       ylim=c(0.5,3.0), ylab="",

       smooth=0.4,

       yolabs=seq(-1,3,by=0.5),

       xlab="proportion of the original cohort")

title(main=paste("(b) entropy = ",format(entropy(g1A,g10),digits=3)),cex=0.6)

abline(h=1,lty=2)

gA0 <- reldist(y=recent$chpermwage, yo=original$chpermwage,

       ywgt=recent$wgt, yowgt=original$wgt,

       smooth=0.4,

       matchlocation=T,

       bar=T,

       ylim=c(0.5,3.0), ylab="",

       yolabs=seq(-1,3,by=0.5),

       xlab="proportion of the original cohort")

title(main=paste("(c) entropy = ",format(gA0$entropy,digits=3)),cex=0.6)
```

```
abline(h=1,lty=2)
```

Panel (a) shows the overall relative density (and is the same as Figure 2 (b)). Panel (b) represents the effect of the median shift in the wage gains between the two cohorts – displaying what the relative density would have looked like if there had been no change in distributional shape. The option `matchshape=T` is used to additively shift the reference sample median to the comparison sample median before comparing the two distributions. Panel (c) displays the effect of changes in distributional shape. The option `matchlocation=T` is used to additively scale the reference sample to the comparison sample before comparing the two distributions.

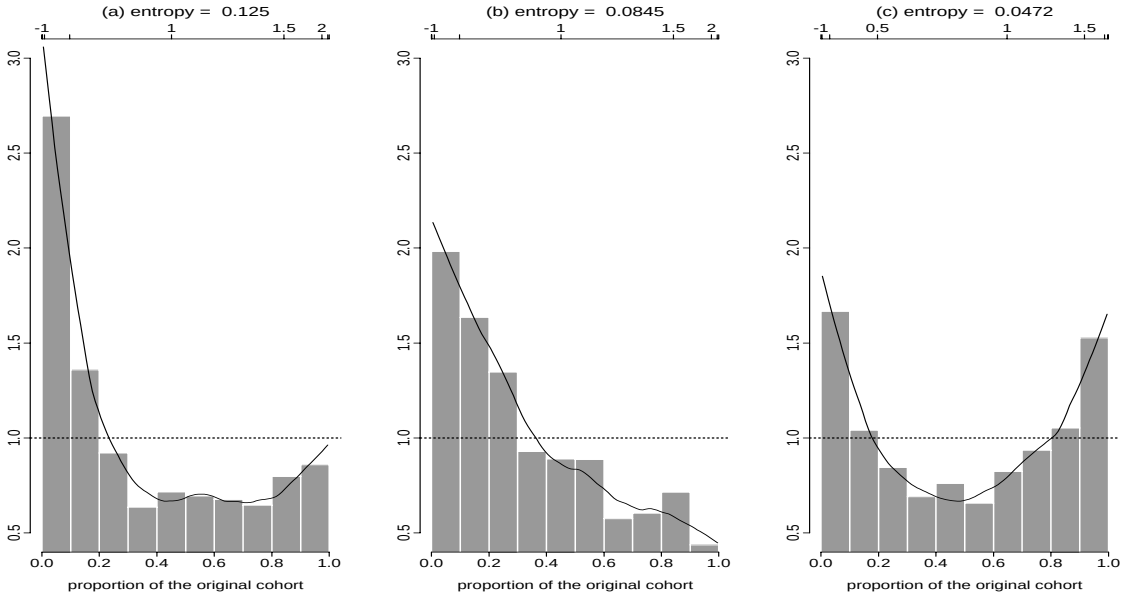

Figure 4. Decomposing the relative distribution of permanent wage growth in the recent and original NLS cohorts into the impact of changes in medians and changes in shape. (a) The (unadjusted) relative density of wage growth; (b) the effect of the median difference in wage growth between the cohorts; (c) the median-adjusted relative density of wage growth (the effect of changes in distributional shape).

5. SUMMARY MEASURES

To complement the graphical displays of the preceding sections, we compute summary measures based on the relative distribution which can be used for the comparison of distributional change. In particular, we calculate entropy, which is a widely used measure of the dispersion of the distribution, and the median relative polarization index, which provides a means to measure distributional polarization. Both summary measures have useful decompositions. The overall entropy may be decomposed into a median effect and a shape effect. The median relative polarization index may be decomposed into upper and lower polarization indices, representing the contributions made by components above and below the median of the relative distribution, respectively. In Table 1 the full set of summary statistics is presented. Note also that entropy summaries are given on the top of Figure 4.

The summary statistics may be reproduced with

```
format(rpy(y=recent$chpermwage,yo=original$chpermwage,
       ywgt=recent$wgt,yowgt=original$wgt,pvalue=T),
       digits=3)
format(rpluy(y=recent$chpermwage,yo=original$chpermwage,
       ywgt=recent$wgt,yowgt=original$wgt,pvalue=T),
       digits=3)
format(rpluy(y=recent$chpermwage,yo=original$chpermwage,
       ywgt=recent$wgt,yowgt=original$wgt,upper=T,pvalue=T),
       digits=3)
```

TABLE 2.

Summary Statistics for the Location/Shape Decomposition of the Relative Distribution of Wage Gains: Recent to Original NLS Cohort

Entropy	Estimate		
overall change in wage growth	0.125		
median effect	0.078		
shape effect	0.047		
percent due to median	62.4%		
percent due to shape	37.6		
Polarization Index	Estimate	95% CI	p–value
Median Index	0.183	0.148 – 0.219	0.000
Lower Index	0.190	0.118 – 0.262	0.000
Upper Index	0.176	0.104 – 0.249	0.000

6. COVARIATE ADJUSTMENT

One can separate the impacts of changes in population composition from changes in the covariate-outcome relationship by adjusting the relative distribution for changes in the distribution of other covariates. This method decomposes the relative distribution into the composition effect or the component that represents the effect of changes in the marginal distribution of the covariate, and a component that represents residual changes.

As educational composition of the NLS cohorts may have changed, the covariate adjustment technique can be used to determine whether differences in the educational profile between the two cohorts explain some of the changes in relative wage gains. Figure 5 shows the relative distribution of final observed education in the two cohorts and is generated by

```
e1 <- original$endeduc
```

```
e1[e1 < 8] <- 8

e1[e1 > 18] <- 18

e2 <- recent$endeduc

e2[e2 < 8] <- 8

e2[e2 > 18] <- 18

postscript("Fig4.ps", width=4.5,height=4.5,horiz=F)

par(err=-1)

g10 <- rddist(y=e2, yo=e1, pool=1, ci=F,

       ywgt=recent$wgt,yowgt=original$wgt,

       yolabs=sort(unique(e1)),

       ylab="relative density",

       xlab="proportion of the original cohort")

title(sub=paste("entropy = ",format(entropy(g10),digits=3)))

abline(h=1,lty=2)
```

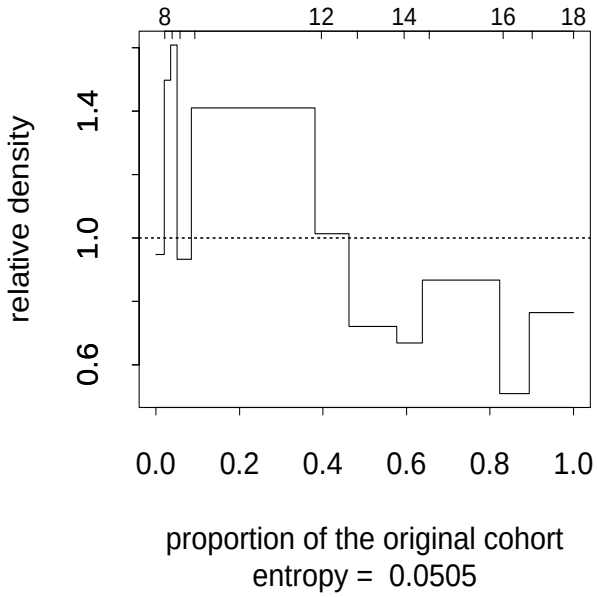

Figure 5. The relative distribution of education for the recent to the original cohort. The upper axis indicates the final number of years of schooling completed.

Figure 6 is a graphical representation of the adjustment of the relative distribution for education composition changes and is produced with

```
par(err=-1)

par(mfrow=c(1,3))
```

```
i3x <- sample(seq(along=original$chpermwage),
       size = 10*length(original$chpermwage),
       prob=rdsamp(e2,e1,recent$wgt,original$wgt),
       replace = T)
schpermwage1 <- original$chpermwage[i3x]
wschpermwage1 <- original$wgt[i3x]
g10 <- reldist(y=recent$chpermwage, yo=original$chpermwage, smooth=0.4, ci=F,
       ywgt=recent$wgt, yowgt=original$wgt,
       yolabs=seq(-1,3,by=0.5), ylim=c(0.5,3.0),
       bar=T,
       xlab="proportion of the original cohort")
title(main=paste("(a) entropy = ",format(g10$entropy,digits=3)),cex=0.6)
abline(h=1,lty=2)
g1A <- reldist(y=schpermwage1, yo=original$chpermwage,
       yowgt=original$wgt, ywgt=wschpermwage1,
       bar=T,
       ylim=c(0.5,3.0), ylab="",
       smooth=0.4, ci=F,
       yolabs=seq(-1,3,by=0.5),
       xlab="proportion of the original cohort")
title(main=paste("(b) entropy = ",format(entropy(g1A,g10),digits=3)),cex=0.6)
abline(h=1,lty=2)
gA0 <- reldist(y=recent$chpermwage, yo=schpermwage1, smooth=0.4, ci=F,
       ywgt=recent$wgt, yowgt=wschpermwage1,
       bar=T,
       ylim=c(0.5,3.0), ylab="",
       yolabs=seq(-1,3,by=0.5),
       xlab="proportion of the original cohort")
title(main=paste("(c) entropy = ",format(gA0$entropy,digits=3)),cex=0.6)
abline(h=1,lty=2)
```

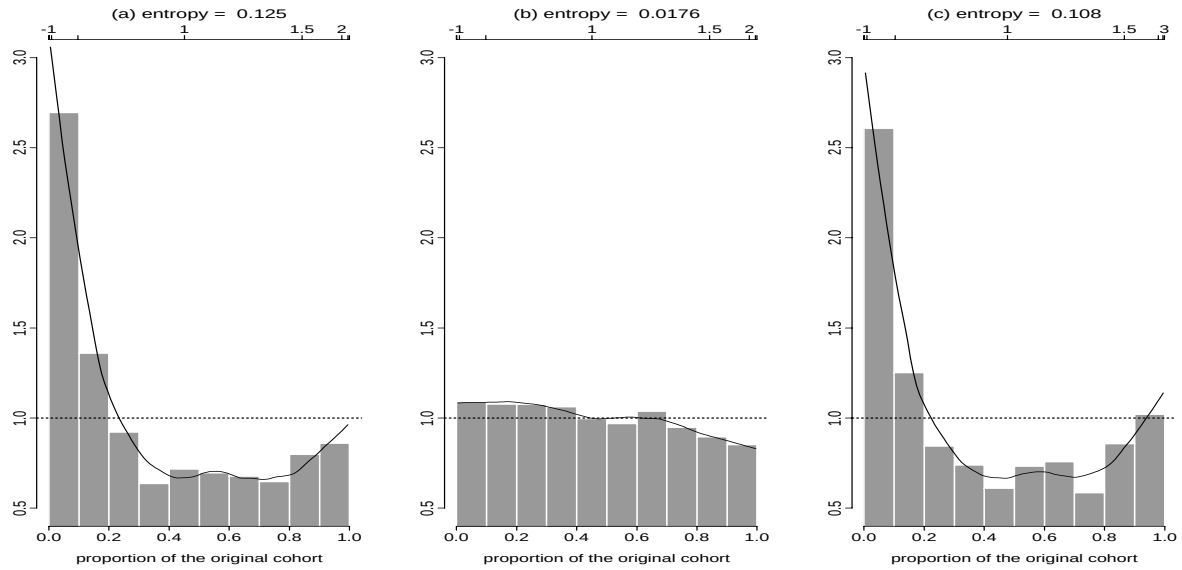

Figure 6. Adjusting the relative distribution of permanent wage growth for changes in the education composition between the two cohorts. (a) The (unadjusted) relative density of wage growth; (b) the effect of changes in the education profile between the cohorts; (c) the education-adjusted relative density of wage growth.

Panel (a) is the (unadjusted) relative density of wage gains (same as Figure 2b), panel (b) represents the education composition effects, and panel (c) represents the *education-adjusted* relative density of wage gains. Thus panel (c) represents the expected relative density of wage gains had the education profiles of the two cohorts been identical.

7. ADDITIONAL TOPICS

For a discrete covariate, we may adjust for this covariate as in Section 6, or we may compare the groups defined by the covariate directly. To demonstrate this technique, education is again used as a covariate, but now it is defined in discrete form. In particular education is divided into the categories of those with a high school degree or less and those with one or more years of college.

Figure 7 compares the distributions of wage gains for the two education groups, as density overlays (a and c) and as relative densities, recent to original cohort (b and d). Panels (a) and (b) compare the wage gains for the high school educated across the two cohorts. Panels (c) and (d) compare the wage gains for the the college educated across the two cohorts. The plots are generated by

```
par(err=-1)
par(mfrow=c(2,2))
spwhso<-sample(pwhso,size=(100000),prob=wgthso/sum(wgthso),replace=T)
spwsco<-sample(pwsco,size=(100000),prob=wgtsco/sum(wgtsco),replace=T)
spwhsr<-sample(pwhsr,size=(100000),prob=wgthsr/sum(wgthsr),replace=T)
```

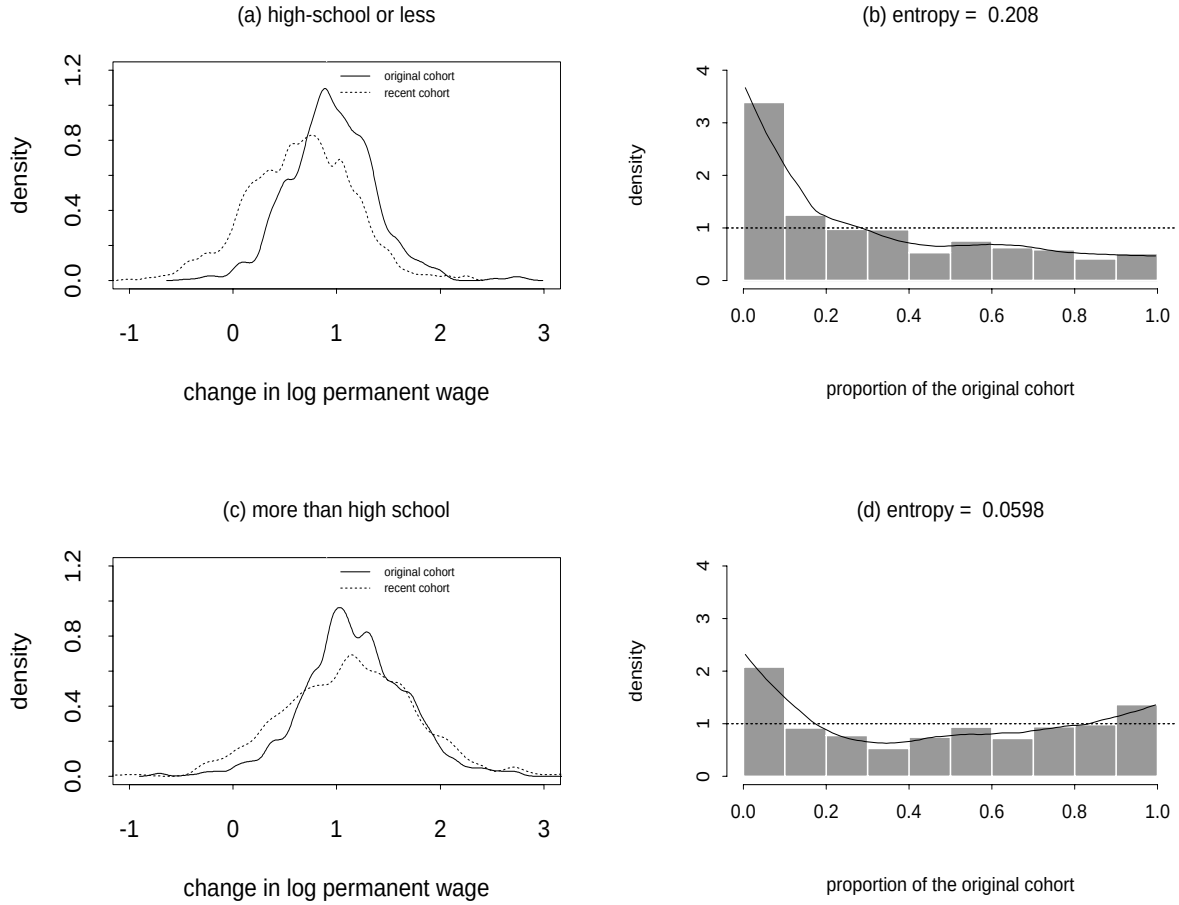

Figure 7. The PDF overlays and cohort relative distributions of permanent wage growth for high school and college-educated workers in the NLS. (a) wage gain PDFs for workers with high school or less education in each cohort; (b) cohort relative distribution (R:O) for those with high school or less; (c) wage gain PDFs for workers with some college in each cohort; (d) cohort relative distribution (R:O) for those with some college. A decile bar chart is superimposed on the relative density estimates.

```
spwscr<-sample(pwscr,size=(100000),prob=wgtscr/sum(wgtscr),replace=T)

nbar <- log(length(spwhso), base = 2) + 1

kwidth <- diff(range(spwhso))/nbar * 0.5

kwidth <- 1.4*kwidth

dens1 <- density(spwhso, n = 500, width=2.0*kwidth)

         plot(x = (dens1$x), y = dens1$y, type = "l",

         xlab = "change in log permanent wage", ylab = "density",

         xlim = c(-1, 3),ylim=c(0,1.2))

fig1legend <- list(x=c(0.9,0.9),y=c(1.25,1.25))

legend(fig1legend,lty=1:2,cex=0.5, bty="n",

  legend=c("original cohort","recent cohort"))

title(main=paste("(a) high-school or less"),cex=0.6)

dens2 <- density(spwhsr, n = 500, width=1.7*kwidth)
```

```
lines(x = (dens2$x), y = dens2$y, type = "l",lty=2)
g10hs <- reldist(y=pwhsr, yo=pwhso, ci=F, smooth=0.4,
         ywgt=wgthsr, yowgt=wgthso,
         bar=T,
         ylim=c(0,4),
         xlab="proportion of the original cohort")
title(main=paste("(b) entropy = ",format(g10hs$entropy,digits=3)),cex=0.6)
abline(h=1,lty=2)
nbar <- log(length(spwscr), base = 2) + 1
kwidth <- diff(range(spwscr))/nbar * 0.5
kwidth <- 1.2*kwidth
dens1 <- density(spwsco, n = 500, width=1.5*kwidth)
plot(x = (dens1$x), y = dens1$y, type = "l",
  xlab = "change in log permanent wage", ylab = "density",
  xlim = c(-1, 3),ylim=c(0,1.2))
fig1legend <- list(x=c(0.9,0.9),y=c(1.25,1.25))
legend(fig1legend,lty=1:2,cex=0.5, bty="n",
  legend=c("original cohort","recent cohort"))
title(main=paste("(c) more than high school"),cex=0.6)
dens2 <- density(spwscr, n = 500, width=2*kwidth)
lines(x = (dens2$x), y = dens2$y, type = "l",lty=2)
g10sc <- reldist(y=pwscr, yo=pwsco, ci=F, smooth=0.4,
         ywgt=wgtscr, yowgt=wgtsco,
         bar=T,
         ylim=c(0,4),
         xlab="proportion of the original cohort")
title(main=paste("(d) entropy = ",format(g10sc$entropy,digits=3)),cex=0.6)
abline(h=1,lty=2)
```

To assess how much the location and shape shifts in each groups' distribution contributes to the overall change in their relative positions, we make a decomposition into the "marginal effects" of each change. It is also possible to obtain a unique decomposition by defining the effects sequentially. Figure 8 presents the two compositions side by side and is produced with

```
par(err=-1)
par(mfrow=c(1,2))
rdhsrscr <- deciles(y=pwhsr, yo=pwscr, ywgt=wgthsr, yowgt=wgtscr,
            binn=binn)
rdhsosco <- deciles(y=pwhso, yo=pwsco, ywgt=wgthso, yowgt=wgtsco,
            binn=binn)
mscrdhsrscr <- deciles(y=pwhsr - wtd.median(pwhsr, weight=wgthsr) +
               wtd.median(pwhso, weight=wgthso), yo=pwscr -
               wtd.median(pwscr, weight=wgtscr) +
               wtd.median(pwsco, weight=wgtsco),
               ywgt=wgthsr, yowgt=wgtscr, binn=binn)
mhsrdhsrscr <- deciles(y=pwhso - wtd.median(pwhso, weight=wgthso) +
               wtd.median(pwhsr, weight=wgthsr), yo=pwsco -
               wtd.median(pwsco, weight=wgtsco) +
               wtd.median(pwscr, weight=wgtscr),
               ywgt=wgthso, yowgt=wgtsco, binn=binn)
m1rdhsrscr <- deciles(yo=pwsco, y=pwhsr - wtd.median(pwhsr, weight=wgthsr)
              + wtd.median(pwhso, weight=wgthso),
              yowgt=wgtsco, ywgt=wgthsr, binn=binn)
m2rdhsrscr <- deciles(y=pwhso, yo=pwscr - wtd.median(pwscr, weight=wgtscr)
              + wtd.median(pwsco, weight=wgtsco),
              ywgt=wgthso, yowgt=wgtscr, binn=binn)
m3rdhsrscr <- deciles(y=pwhsr, yo=pwsco - wtd.median(pwsco, weight=wgtsco)
              + wtd.median(pwscr, weight=wgtscr),
              yowgt=wgtsco, ywgt=wgtscr, binn=binn)
```

```
achange <- binn*(rdhsrscr$x - rdhsosco$x)
armeff  <- binn*(mhsrdhsrscr$x - rdhsosco$x)
ahseff  <- binn*(m1rdhsrscr$x - rdhsosco$x)
asceff  <- binn*(m2rdhsrscr$x - rdhsosco$x)
ainteff <- achange - armeff - ahseff - asceff
barplot(height=achange,histo=T,width=(1:binn)-0.5,axes=F,
  xlab="Decile",ylab="Percentage Point Change", ylim=c(-20.0,25))
axis(1,labels=T,at=(1:binn))
axis(2,labels=T,at=seq(-20.0,25,length=10))
title(main="(a) Marginal effects",cex=0.6)
lines(y=(armeff),x=(1:binn),lty=1)
lines(y=(asceff),x=(1:binn),lty=3)
lines(y=(ahseff),x=(1:binn),lty=2)
abline(h=seq(-20,25,length=10),lty=2)
points(y=(armeff),x=(1:binn),mark=16,cex=0.7)
points(y=(asceff),x=(1:binn),mark=3,cex=0.7)
points(y=(ahseff),x=(1:binn),mark=1,cex=0.7)
fig1legend <- list(x=c(4,4),y=c(25,25))
legend(fig1legend,mark=c(16,1,3),lty=c(1:3),cex=0.5, bty="n",
legend=c("Change in relative median",
  "High-school shape effect","College shape effect"))
armeff  <- binn*(mhsrdhsrscr$x - rdhsosco$x)
ahseff  <- binn*(m3rdhsrscr$x - mhsrdhsrscr$x)
asceff  <- binn*(rdhsrscr$x - m3rdhsrscr$x)
barplot(height=achange,histo=T,width=(1:binn)-0.5,axes=F,
  xlab="Decile",ylab="Percentage Point Change",
  ylim=c(-20.0,25))
axis(1,labels=T,at=(1:binn))
axis(2,labels=T,at=seq(-20.0,25,length=10))
title(main="(b) Sequential effects",cex=0.6)
```

```
lines(y=(armeff),x=(1:binn),lty=1)

lines(y=(asceff),x=(1:binn),lty=3)

lines(y=(ahseff),x=(1:binn),lty=2)

abline(h=seq(-20,25,length=10),lty=2)

points(y=(armeff),x=(1:binn),mark=16,cex=0.7)

points(y=(asceff),x=(1:binn),mark=3,cex=0.7)

points(y=(ahseff),x=(1:binn),mark=1,cex=0.7)

fig1legend <- list(x=c(4,4),y=c(20,25))

legend(fig1legend,mark=c(16,1,3),lty=c(1:3),cex=0.5, bty="n",

  legend=c("Change in relative median",

  "High-school shape effect","College shape effect"))
```

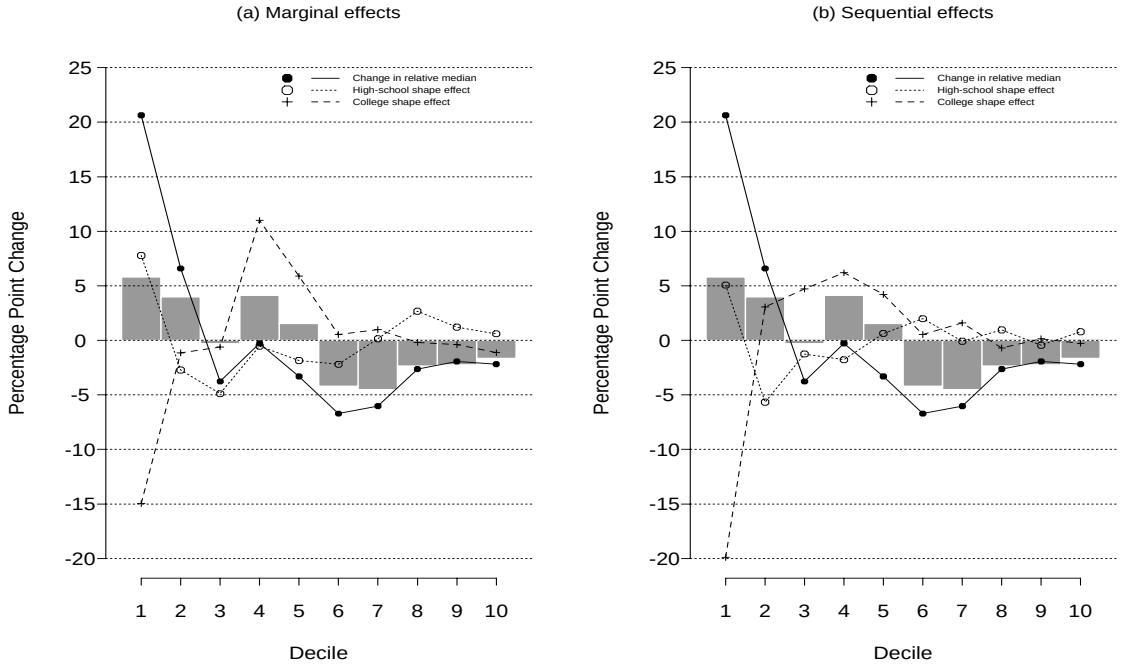

Figure 8. Sources of the change in the cohort relative distribution of wage gains by education level. (a) Marginal effects. (b) Sequential effects.

Panel (a) represents the marginal effects of the median shift from the original density, the marginal effect of the shape change in the high school distribution, and the marginal effect of the shape change in the college distribution. Panel (b) represents the sequential effects of the relative median shift from the original relative distribution, then the shape change in the college distribution form the median shifted original relative distribution, and finally the shape change in the high school distribution from the median shifted, college shape changed relative distribution.

REFERENCES

Handcock, Mark. S., and Morris, Martina (1998). Relative Distribution Methods. *Sociological Methodology,* Vol 28, p. 53-97.