THE IMPACT OF 2000 CENSUS BASED POPULATION CONTROLS ON HEALTH ESTIMATES IN THE NATIONAL HEALTH INTERVIEW SURVEY

Carrie Lynch and Van Parsons, National Center for Health Statistics (NCHS)
Carrie Lynch, NCHS, 3311 Toledo Road, Room 3118, Hyattsville, MD 20782

KEY WORDS: multiple-race, population controls, weights, standard errors

1. Introduction

The Office of Management and Budget (OMB) issued guidelines for the collection and presentation of Federal data on race and ethnicity in the 1977 Statistical Policy Directive No. 15 (OMB, 1977). In response to the nation's population becoming more diversified as a result of increasing immigration and interracial marriages, the OMB revised the standards in 1997 (OMB, 1997). The U.S. Bureau of the Census adopted these new guidelines in the 2000 Census, and other U.S. Federal statistical programs were required to comply by January 1, 2003.

One of the major revisions in the directive has allowed individuals to select one or more race categories when responding to Federal surveys, administrative forms and records, and other data collections. The four single-race categories established in the 1977 standards were White, Black, Asian or Pacific Islander (API), and American Indian or Alaska Native (AIAN). In the 1997 standards, the number of single-race categories increased to five: White, Black or African American, Asian, Native Hawaiian or Other Pacific Islander (NHOPI), and AIAN. The introduction of multiple-race reporting yielded thirty-one race categories.

The National Health Interview Survey (NHIS), which is a major national household survey that monitors the health of the U.S. civilian noninstitutionalized population (CNIP), has allowed respondents to select more than one race category since 1976. The interviewer shows the respondent a card with numbered race categories and asks, "What race do you consider yourself to be? Please select one or more of these categories." If the respondent selects more than one race category, the interviewer asks a follow-up question to obtain a 'primary' race, "Which one of these groups, that is (the interviewer reads the groups) would you say best represents your race?"

Although the NHIS has been following many aspects of the new standards for many years, it became fully compliant with the fielding of the 1999 NHIS (National Center for Health Statistics, 2002). However, the sample weights produced using 2000 Census based population controls will not be used until the release of the 2003 NHIS public use file. Public use files for 2000, 2001, and 2002 use 1990-based sample weights.

This paper discusses the impact of 2000 Census based population controls on health estimates in the 2002 NHIS. Section 2 contains procedures for bridging data collected under the 1977 and 1997 standards and criteria for evaluating the technical merits of the bridging methods. This section also describes the bridging method employed in the development of NHIS 2000-based sample weights. Section 3 discusses the population controls used to weight NHIS data and the change in the controls from the two most recent Decennial Censuses. Section 4 compares selected health estimates using 1990 and 2000 Census based sample weights, and Section 5 concludes with a summary.

2. Bridging

For the development of sample weights, in particular, the poststratification factor, the NHIS data are classified into old single-race categories of Black and all other races (i.e., White, API, and AIAN). On the other hand, the 2000 Census based population controls used to weight sample data are grouped by new single-race and multiple-race categories. Therefore, a method to bridge 2000-based population to old single-race categories is necessary for the development of 2000-based sample weights.

In 2000, the OMB issued a guidance document containing methods to bridge or compare data collected under the 1977 and 1997 standards (OMB, 2000). The guidance document also contains criteria that could be used to evaluate the technical merits of the bridging methods.

2.1 Methods for Bridging

The four whole allocation methods assign each respondent who selects more than one race group to exactly one of the single-race groups marked: the *Smallest Group* assigns multiple-race responses to the single-race group with the smallest population; the *Largest Group* assigns multiple-race responses to the single-race group with the largest population; the *Largest Group Other Than White* assigns multiple-race responses to the Non-White single-race group with the largest population; and the *Plurality* assigns multiple-race responses to the single-race group most frequently identified as the primary race by multiple-

race respondents in the NHIS from the given multiple-race group.

The two fractional allocation methods assign part of each multiple-race response to each of the single-race groups marked: the *Equal Fractions* assigns multiple-race responses equally among the single-race groups; and the *NHIS Fractions* proportionally assigns multiple-race responses among the single-race groups where the weight is derived using responses to the primary race question in the NHIS.

The guidance document did not discuss bridging methods for single-race responses because individuals who select one race under the new standards were assumed to respond similarly under the old standards.

Schenker and Parker (2003) developed another bridging method. The NHIS Regression method is an extension of the NHIS Fractions method and involves estimating logistic regressions from NHIS data to predict the primary race given the multiple-race responses and covariates.

2.2 Evaluation Criteria

There are nine criteria that could be used to evaluate the technical merits of the bridging methods: measure change over time; minimize disruptions to the single-race distribution; have a wide range of applicability; meet confidentiality and reliability standards; be statistically defensible; be easy to use; require relatively little statistical knowledge; be understandable and communicable; and be congruent with the respondent's choice.

Schenker and Parker (2003) evaluated the bridging methods based on statistical defensibility. They concluded that the NHIS Fractions and NHIS Regression methods provide better estimates of the proportion of multiple-race respondents who would have selected single-race categories. Moreover, they demonstrated that in some instances the NHIS Regression method produces less biased point estimates and smaller variance estimates compared to the NHIS Fractions method.

For the purpose of computing poststratification factors in the development of sample weights, the NHIS Fractions method was selected as the most appropriate method for bridging Census population controls distributed by multiple-race groups to single-race groups. The NHIS Fractions method was preferred over the other OMB bridging methods because of the recommendation by Schenker and Parker and favored over the NHIS Regression method because it is easier to use, requires less statistical knowledge, and is easier to explain to the public.

2.3 NHIS Fractions Method

Since the NHIS age-sex-race/ethnicity poststratification strata consist of the race/ethnicity categories of Hispanic, Non-Hispanic (NH) Black, and NH-Other, the NHIS Fractions method was only used to bridge NH-Black multiple-race population (e.g., NH-Black, NH-White, and NH-NHOPI) to old single-race categories. All other Non-Hispanic multiple-race population (e.g., NH-Asian and NH-White) was assigned to the NH-Other category, and Hispanic multiple-race population (e.g., Hispanic Black and Hispanic White) was assigned to the Hispanic category.

The 1997-2001 NHIS In-house files used to compute NHIS fractions were restricted to NH-Black multiple-race responses to determine what percentage of them identify most closely with each old single-race category. The data were also limited to persons who answered the Hispanic origin question; persons with nonzero sample weights; and NH-Black multiple-race persons who answered the primary race question. The sample weights were also divided by five since data were combined over survey years.

In addition, some multiple-race responses were recoded. Multiple-race responses of "Asian" or "NHOPI" were assigned to the old racial category of "API." Multiple-race respondents who reported "Other" as one of their multiple-races were treated as if "Other" was not selected. For example, multiple-race respondents who selected "Black" and "Other" were recoded to single-race "Black." Similarly, multiple-race respondents who selected "Asian" "White" "Black" and "Other" were reclassified as "Asian" "White" and "Black." Furthermore, the fifteen standard primary race responses were reclassified as either "Black" or "Non-Black." The remaining primary race responses were deleted from the analysis.

Table 1 shows that there were 2,080 NH-Black multiple-race individuals in the NHIS during 1997-2001. Of these, 79 percent selected a primary race. The percent by each multiple-race category ranged from 68 percent (API/Black/White) to 92 percent (AIAN/Black) for categories with at least thirty respondents. Table 1 also shows the largest NH-Black multiple-race category was Black/White with 1,292 respondents or 62 percent.

Table 2 contains the NHIS fractions or weighted proportions of NH-Black multiple-race respondents by primary race and sex-age. It can be seen that NH-Black multiple-race respondents identify themselves as Black more than all other race groups combined. The sex-age categories and the one all inclusive multiple-race category presented in Table 2 were selected because they satisfied the sample size

requirement of at least thirty respondents and minimized the variability in the proportions.

The new population base controls were allocated to the race/ethnicity categories of Hispanic, NH-Black, and NH-Other (referred to as Non-Black as described above) within each sex-age category, as follows: proportionally assigned NH-Black multiple-race population to both Black and Non-Black categories using the NHIS fractions contained in Table 2; assigned all other multiple-race population, not of Hispanic origin, to the Non-Black category; and assigned Hispanic multiple-race population to the Hispanic category.

3. Population Controls

Every month the U.S. Bureau of the Census develops CNIP controls for the purpose of second stage weighting of the Current Population Survey data (Bureau of Labor Statistics et al, 2002). Historically, the NHIS, like other population-based surveys, uses similar CNIP controls to weight sample data and ensure that national population estimates distributed by age-sex-race/ethnicity categories from the two surveys agree.

3.1 Using Population Controls to Create Weights

Weights are created for the NHIS sample every quarter and are divided by four to create annual weights which are used when a calendar year of data become available. The quarterly weights allow for national estimates to be made for each calendar quarter of sample data.

The national person weight is the product of four factors: inverse of the probability of selection; household nonresponse adjustment; first-stage ratio adjustment; and second-stage ratio adjustment. The fourth factor, commonly referred to as the poststratification factor, is the only component that employs the quarterly population controls, and hence requires recalculation to determine the impact of new population controls on health estimates. The poststratification factor is equal to the population control for that poststratum divided by the sum of the weights prior to the application of the second-stage ratio adjustment for respondents in the poststratum.

The sample adult and sample child weights are derived by dividing the person weight by the poststratification factor, adjusting for additional levels of sampling, and applying a new poststratification adjustment, which has fewer poststratification strata. Refer to the design and estimation report (Botman et al, 2000) and survey documentation (National Center for Health Statistics, 2002) for more information about weights.

3.2 Change in Population Controls

The effects of the new population base controls on the analysis of 2002 NHIS data were examined using two sets of 2002 quarterly (February 1, May 1, August 1, November 1) population controls. The first set reflected the results of the 1990 Census, and the second set used 2000 Census information. When comparing the two sets of controls, it is important to know that the 1990-based controls include an adjustment for the net Census under enumeration, whereas the 2000-based controls do not include an adjustment. The U.S. Bureau of the Census asserts that the differences among adjusted and unadjusted counts from the 2000 Census are very small as compared to the corresponding differences in the 1990 Census, and thus, recommends the use of unadjusted 2000-based controls.

Switching to the 2000-based controls increased the CNIP by about 1.2 percent (3.5 million) in the fourth quarter of 2002. The population levels rose about 7.3 percent for Hispanic persons, 1.2 percent for NH-Black persons, and 0.2 percent for NH-Other persons. The percent changes in the other three quarters were very similar to the fourth quarter, so they are not discussed.

Table 3 contains the percent changes in population controls by the NHIS person level poststratification strata. The most striking changes occurred in the Hispanic population. The number of Hispanic males 25-29 years of age increased about 41 percent. There were also large percent increases for Hispanic males ages 20-24 years (21 percent) and ages 30-34 years (31 percent). The population of Hispanic females 20-24 years of age increased almost 7 percent, and the percent increases for Hispanic females ages 25-29 years and 30-34 years were between 18 and 19 percent. The number of Hispanic persons ages 65 years and over decreased approximately 11 percent for males and 10 percent for females.

Some changes were also noted for the other two race/ethnicity groups. The new population controls increased the population of NH-Black females ages 75 years and over (8 percent) and females ages 5-9 years (6 percent). The population of NH-Black males ages 50-54 years and 55-64 years rose about 7 percent and males ages 5-9 years increased almost 6 percent. The changes in NH-Other population sizes by sex-age categories were less than 4 percent, except for females ages 75 years and over where there was a 6 percent increase.

The large percent changes highlighted in the two preceding paragraphs indicate that the weights for these domains will be different using the new 2000-based population controls. Since the switch will increase (or decrease depending on the domain) the

number of persons reporting and not reporting a health condition proportionately, one should expect less of an impact on the weighted estimate of a proportion than on an estimated total. On the other hand, without the investigation of the sampling error, it is uncertain whether the population changes will yield significant differences in first-order estimates for these domains.

4. Impact on Health Estimates

The data for this research was obtained from the Person, Sample Adult, and Sample Child 2002 NHIS Early Release in-house files because final data were not available. Early Release data are subject to change during final processing. However, it is believed that the general conclusions drawn from this research would be the same if final data were used. Table 4 lists the health conditions extracted from each file.

The domains of study were five race/ethnicity groups (total, Hispanic, NH-White, NH-Black, NH-Other) x three sex groups (total, male, female) x age groups. There were seven age groups (total, under 5 years, 5-17 years, 18-44 years, 45-64 years, 65-74 years, 75 years and over) for person health conditions, five age groups (total, 18-44 years, 45-64 years, 65-74 years, 75 years and over) for sample adult health conditions, and three age groups (total, under 5 years, 5-17 years) for sample child health conditions.

For sample sizes of at least thirty persons, hypothesis tests were conducted to determine whether the proportions of persons reporting health conditions using the 1990-based sample weights were statistically different than the proportions using the 2000-based sample weights. The estimates of the difference of proportions (DIFF) and standard error of the difference of proportions (seDIFF) for each health condition by domain were calculated using a Taylor series linearization method in SAS-callable SUDAAN software (Research Triangle Institute, 2001). See the appendix for techniques used to approximate the standard errors of the difference of proportions and totals. Significant differences were dependent upon the value of the test statistic, equal to DIFF divided by seDIFF, and the critical value for a two-tailed t test at the 5 percent level of significance. Hypothesis tests were also conducted to detect significant differences in the total number of persons reporting health conditions using the different weights.

Care must be used when interpreting a significant difference. For example, the proportion of NH-White persons ages 65-74 years reporting a hospital stay using 1990-based weights (.16242503) and the proportion using 2000-based weights (.16238498) were significantly different. Since the two compared statistics were highly correlated, the standard error of the difference was small. This fact, in turn, inflated the value of the test statistic (DIFF/seDIFF = -4.0050 x 10^{-5}/1.6158 x 10^{-5} = -2.4786). Comparing the values of the test statistic and the critical value (-1.9605), the compared statistics were significantly different even though their numerical values were the same to four significant digits. Applying statistical methods to make inferences for two population means using a paired sample concluded a significant difference, but in practice the proportions would be the same. **A criterion was developed whereby only significant differences that have any practical value would be considered. The change of at least 1 percent for *proportions* and the percent change of at least 5 percent for *totals* were chosen.** Therefore, in the preceding example, according to the criterion, the two compared statistics would not be considered significantly different.

4.1 Difference of Proportions

Of the 1,196 health estimates (368 person + 738 sample adult + 90 sample child), the *proportion* of persons reporting a health condition using the 1990-based weight was statistically different than the *proportion* using the 2000-based weight for 3 of the health estimates. Table 5 shows that each one came from the Hispanic population.

4.2 Difference of Totals

Table 6 shows the health estimates for which a significant difference in the *total* number of persons reporting a health condition using the 1990-based weight versus the 2000-based weight was detected using health conditions in the Person file. Of the 368 health estimates, significant differences were detected in 13 percent of the health estimates. Of these, 68 percent were Hispanic domains. The new controls primarily affected Hispanic male, female, and total persons ages 18-44 years, 65-74 years, and 75 years and over and the other race/ethnicity groups for females ages 75 years and over.

Tables are not shown for significant differences using health conditions in the Sample Adult and Sample Child files, but their results are discussed next. Examining the *total* number of persons reporting health conditions in the Sample Adult file, significant differences were detected in 16 percent (118 out of 738) of the health estimates. Of these, 85 percent came from the Hispanic population. The only Hispanic estimates not impacted by the new controls were male, female, and total persons 45-64

years of age. The switch in controls affected a few of the health conditions for NH-White, NH-Other, and Total females ages 75 years and over and NH-Black males 45-64 years of age.

Of the 90 health estimates from the Sample Child file, significant differences were detected in 1 percent of the health estimates. The problematic estimate was the *total* number of Hispanic females ages 5-17 years reporting trouble hearing.

5. Summary

The purpose of this research was to determine whether switching from the 1990 to 2000 Census based population controls would affect health estimates in the 2002 NHIS. The switch increased the civilian noninstitutionalized population by about 1.2 percent (3.5 million) in the fourth quarter of 2002. The most striking population size changes occurred in the Hispanic population. These and other notable changes in population levels for other race/ethnicity groups had a direct impact on the sample weights.

The research presented in this paper demonstrated that the change in population controls affected weighted estimates of *totals* more than *proportions*. The switch impacted Hispanic estimates more than the other race/ethnicity groups because the Hispanic population had the largest rate of growth between the 1990 and 2000 Censuses. The majority of the affected age groups were 18-44 years, 65-74 years, and 75 years and over for Hispanics and 75 years and over for the other race/ethnicity groups.

Users of NHIS data should use caution when comparing estimates, particularly of counts, from survey years which use different population based controls to compute weights. Any detected significant differences in persons reporting health conditions may not be due to change in prevalence but may be due to controls used to compute weights.

References

Botman, S., Moore, T., Moriarity, C., and Parsons, V. (2000), "Design and Estimation for the National Health Interview Survey, 1995-2004," National Center for Health Statistics, *Vital Health Stat*, 2(130).

Bureau of Labor Statistics, Economics and Statistics Administration, and U.S. Census Bureau (2002), "Current Population Survey, Design and Methodology," Technical Paper 63RV, Washington, DC: Government Printing Office.

National Center for Health Statistics (2002), "2000 National Health Interview Survey (NHIS) Public Use Data Release, NHIS Survey Description."

Office of Management and Budget (2000), "Provisional Guidance on the Implementation of the 1997 Standards for the Collection of Federal Data on Race and Ethnicity."

Office of Management and Budget (1977), "Race and Ethnic Standards for Federal Statistics and Administrative Reporting, Statistical Policy Directive 15."

Office of Management and Budget (1997), "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity," Federal Register 62FR58781-58790.

Research Triangle Institute (2001), *SUDAAN User's Manual*, *Release 8.0*, Research Triangle Park, NC: Research Triangle Institute.

Schenker, N. and Parker, J.D. (2003), "From Single-race Reporting to Multiple-race Reporting: Using Imputation Methods to Bridge the Transition," *Statistics In Medicine*, 22, 1571-1587.

Acknowledgement

The authors thank Dr. Nathaniel Schenker of the Office of the Director, National Center for Health Statistics, for providing consultation on the NHIS Fractions method.

Table 1. Sample Size and Number and Percent of Primary Race Selections for Non-Hispanic Black Multiple-Race Respondents, 1997-2001 NHIS In-house Files

Multiple-Race	Sample Size	Number of Respondents Selected Primary Race	Percent of Respondents Selected Primary Race
AIAN[1]/Black	430	397	92
API[2]/Black	136	116	85
Black/White	1,292	949	73
AIAN[1]/API[2]/Black	12	9	75
AIAN[1]/Black/White	167	136	81
API[2]/Black/White	40	27	68
AIAN[1]/API[2]/Black/White	3	1	33
Total	2,080	1,635	79

[1] American Indian and Alaska Native
[2] Asian and Pacific Islander

Table 2. Sample Size and NHIS Fractions for Non-Hispanic Black Multiple-Race Respondents by Sex and Age, 1997-2001 NHIS In-house Files

Sex and Age	Sample Size	NHIS Fraction Black	NHIS Fraction Non-Black
Female			
Under 18 years	489	.61	.39
18-44 years	251	.74	.26
45 years and over	139	.79	.21
Male			
Under 18 years	469	.67	.33
18-44 years	192	.76	.24
45 years and over	95	.80	.20
Total	1,635		

Table 3. Percent Change in Population Controls by NHIS Person Level Poststratification Strata when Switching from 1990 Census Base to 2000 Census Base, 2002 Quarter 4

Age	Female			Male		
	Hispanic	NH-Black	NH-Other	Hispanic	NH-Black	NH-Other
Under 1 year	4.0	2.0	-0.6	4.1	4.0	-0.5
1-4 years	1.7	3.0	-2.8	2.0	3.7	-2.8
5-9 years	3.0	5.7	-3.2	3.3	6.2	-3.0
10-14 years	5.9	1.6	-2.0	6.0	1.8	-1.8
15-17 years	4.7	-1.5	-1.1	2.5	-2.6	-1.6
18-19 years	-0.1	-2.5	-3.0	2.1	1.4	-2.5
20-24 years	7.4	0.5	-0.8	20.9	2.1	-3.5
25-29 years	18.5	-1.2	0.6	40.8	-1.0	-0.2
30-34 years	18.1	1.2	1.1	30.6	-2.1	2.0
35-44 years	4.9	0.3	-0.4	8.2	-4.2	-1.5
45-49 years	1.2	0.4	0.6	2.5	-1.2	-0.1
50-54 years	1.5	3.9	3.0	3.8	6.5	3.0
55-64 years	-3.0	1.8	2.9	-0.2	6.8	3.0
65-74 years[1]		0.8	3.2		-2.8	2.7
75 years and over[1]	-9.5	7.9	6.0	-10.9	-4.4	1.1

[1] The ages 65-74 years and 75 years and over were collapsed into 65 years and over for Hispanic persons only.

Table 4. Health Conditions Assessed

NHIS File	Health Condition	
Person	▪ Health insurance ▪ Activities of daily living	▪ Hospital stay ▪ Health status
Sample Adult	▪ Current smoker ▪ Ever smoked ▪ Diabetes ▪ Hearing ▪ Cancer	▪ Hypertension ▪ Overweight ▪ Obesity ▪ Heart disease ▪ Asthma
Sample Child	▪ Hearing	▪ Asthma

Table 5. Significant Differences in Proportions, 2002 NHIS Early Release Files

NHIS File and Health Condition	Race/Ethnicity	Sex	Age
▪Person File Health Insurance	Hispanic	Male	Total
▪Sample Adult File Hypertension Hypertension	Hispanic Hispanic	Total Male	Total Total
▪ Sample Child File ----[1]	----[1]	----[1]	----[1]

[1] No entries

Table 6. Significant Differences in Totals, 2002 NHIS Early Release Person File

Age and Sex	Total				Hispanic				NH-White				NH-Black				NH-Other			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Total								X												
Under 5 years	N				N				N				N				N			
5-17 years					N								N				N			
18-44 years					X	X	X	X									N			
45-64 years																	N			
65-74 years					N	X	X	X									N			
75 years and over					X	X	X	X					X				N			
Male								X									N			
Under 5 years	N				N				N				N				N			
5-17 years					N								N				N			
18-44 years					N	X	X	X					N				N			
45-64 years					N								N				N			
65-74 years					N	X	X	X					N				N			
75 years and over					N	X	X	X					N				N			
Female																	N			
Under 5 years	N				N				N				N				N			
5-17 years					N				N				N				N			
18-44 years					N	X	X	X					N				N			
45-64 years																	N			
65-74 years					N	X	X	X					N				N			
75 years and over		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	N	X	X	X

Notes: One (1) represents activities of daily living, 2 represents hospital stay, 3 represents health status, 4 represents health insurance, X denotes a significant difference was detected, and N denotes a hypothesis test was not performed because the sample size was fewer than 30 respondents.

Appendix. Approximation of Standard Errors

The SUDAAN software cannot directly compare the estimates using the 1990-based and 2000-based weights. The following techniques were used to approximate the standard errors of the difference of proportions and totals.

Difference of Proportions

Equation 1 formulates the difference of proportions of persons reporting a health condition using two different weights. Let D denote the domain of study; x_i denote the indicator variable of the health condition for the ith sampling unit; w_{i90} denote the 1990-based sampling weight for the ith sampling unit; and w_{i00} denote the 2000-based sampling weight for the ith sampling unit. The new indicator variable created to estimate the difference of proportions of persons reporting a health condition was defined by equation 2, and the weight variable was w_{i00}. These y_i values were run through SUDAAN using weights w_{i00}, and the standard error was evaluated.

Equation 1:

$$DIFF_D = \frac{\sum_D x_i w_{i00}}{\sum_D w_{i00}} - \frac{\sum_D x_i w_{i90}}{\sum_D w_{i90}} = \sum_D x_i \left[\frac{w_{i00}}{\sum_D w_{i00}} - \frac{w_{i90}}{\sum_D w_{i90}} \right] \cdot \frac{1}{w_{i00}} \cdot w_{i00} .$$

Equation 2:

$$y_i = x_i \left[\frac{w_{i00}}{\sum_D w_{i00}} - \frac{w_{i90}}{\sum_D w_{i90}} \right] \cdot \frac{1}{w_{i00}} .$$

Difference of Totals

Equation 3 formulates the difference of total number of persons reporting a health condition using two different weights. The new indicator variable created to estimate the difference of total numbers of persons reporting a health condition was defined by equation 4, and the weight variable was w_{i00}. These y_i values were run through SUDAAN using weights w_{i00}, and the standard error was evaluated.

Equation 3:

$$DIFF_D = \sum_D x_i w_{i00} - \sum_D x_i w_{i90} = \sum_D x_i \left(\frac{w_{i00} - w_{i90}}{w_{i00}} \right) \cdot w_{i00} .$$

Equation 4:

$$y_i = x_i \left(\frac{w_{i00} - w_{i90}}{w_{i00}} \right) .$$