Reweighting Household Surveys for Tax Microsimulation Modelling: An Application to the New Zealand Household Economic Survey

John Creedy
The University of Melbourne

Ivan Tuckwell
New Zealand Treasury

Abstract
This paper reports a reweighting exercise for the New Zealand Household Economic Survey, which is the basis of the Treasury's microsimulation model, TaxMod. Comparisons of benefit expenditures in a variety of demographic groups, along with population data, reveal that TaxMod estimates differ substantially from totals based on administrative data, when the weights provided by Statistics New Zealand are used. After describing the method used to compute new weights, the calibration requirements are reported. These relate to the age structure of the population and the number of beneficiaries for Unemployment Benefit, Domestic Purposes Benefit, Invalid's and Sickness Benefits and Family Support and Tax Credits. The revised weights and expenditure estimates are reported and the resulting distribution of income examined. The new weights are found to produce much improved expenditure estimates, while having little effect on the resulting income distribution. The effects of reweighting are demonstrated using a simple policy simulation.

1. Introduction
Tax microsimulation models are based on large-scale cross-sectional survey data. Each individual or household has a sample weight provided by the statistical agency responsible for collecting the data. The weights are used for 'grossing up' from the sample in order to obtain estimates of population values. This applies to aggregates such as income taxation, the number of recipients of a particular social transfer, or the number of people in a particular age group. In addition, the weights are also used in the estimation of measures of population inequality and poverty.

Address for correspondence: John Creedy, Department of Economics, The University of Melbourne, Parkville, Vic 3010. Email: jcreedy@unimelb.edu.au
This paper was initially written at the New Zealand Treasury. The views, opinions, findings, and conclusions or recommendations expressed in this Working Paper are strictly those of the authors. They do not necessarily reflect the views of the New Zealand Treasury. We are grateful to the New Zealand Ministry of Social Development and Inland Revenue officials for providing summary information about social expenditures used in this paper, and Patrick Nolan for collating this information. We have benefited from comments on an earlier version by Guyonne Kalb and two referees.

The typical starting point is to use weights that are inversely related to the probability of selecting the individual in a random sample, with some adjustment for non-response. It has become common for data-gathering agencies, using 'minimal' adjustments, to produce revised weights to ensure that, for example, the estimated population age/gender distributions match population totals obtained from other sources, in particular census data. Such methods are well known among survey statisticians.[1] Until relatively recently, applied economists using survey data have simply taken the weights as given, but it has become necessary for microsimulation modellers to look at them more closely. This is because there is no guarantee that weights calibrated on demographic variables produce appropriate revenue and expenditure totals. This is problematic when using a simulation model to examine the likely costs of a hypothetical reform to the tax and transfer system. Reweighting may also be required when using a dataset that is several years old, so that changes in the structure of the population may be expected to have taken place.

This paper reports a reweighting exercise for the New Zealand Household Economic Survey, which is the basis of the Treasury's direct tax and benefit microsimulation model, TaxMod.[2] The Household Economic Survey examines private households from across New Zealand. It collects expenditure data for the entire household and income data for each individual in the household. Each surveyed household has a sample weight provided by Statistics New Zealand.

Section 2 describes the basic method used to compute new weights. The basic approach follows Deville and Särndal (1992). As this does not appear to be well-known among economists, this section sets out the method in some detail.[3] Section 3 compares the expenditure totals produced by TaxMod, using the Household Economic Survey weights provided by Statistics New Zealand, with administrative data relating to actual expenditures. The revised weights and expenditure estimates are reported in section 4. One problem is that producing new weights based on selected conditions may distort other variables of interest, such as incomes. Hence section 5 examines the distribution of income before and after reweighting. An example of a policy simulation, involving an increase in family payments and income tax rates, using two sets of weights is given in section 6. Brief conclusions are in section 7.

2. The Reweighting Procedure

This section describes the use of extraneous information to specify calibration conditions for reweighting, such that the new weights are as close as possible

[1] A detailed description of calibration and Generalised Regression (GREG) methods used in Belgium is given in Vanderhoeft (2001), which also describes the SPSS based program g-CALIB-S. Bell (2000) describes methods used in the Australian Bureau of Statistics household surveys, involving the SAS software GREGWT. Statistics Sweden uses the SAS software CLAN, described by Andersson and Nordberg (1998) and also used by the Finnish Labour Force Survey.

[2] TaxMod reads in one family at a time, calculates market income, adds income from various government programs (benefits, superannuation, Family Support, Accommodation Supplement) according to eligibility, and calculates tax liability. It can provide output at the personal, family and household level. TaxMod assumes that each individual's labour supply remains fixed when the tax and benefit system changes

[3] For an introductory survey of reweighting, see Creedy (2003).

to the initial or 'design' weights. The method therefore requires a distance function to be specified. Subsection 1 provides a formal statement of the optimisation problem, and subsection 2 examines a convenient class of distance functions. An iterative approach for solving the nonlinear first-order conditions, based on Newton's method, is derived in subsection 3. Several alternative distance functions are described in subsection 4.

The Problem

For each of K individuals in a sample survey, information is available about J variables; these are placed in the vector:[4]

$$x_k = \left[x_{k,1},\ldots,x_{k,J}\right]' \tag{1}$$

For present purposes these vectors contain only the variables of interest for the calibration exercise, rather than all measured variables. Most of the elements of x_k are likely to be 0/1 variables. For example $x_{k,j} = 1$ if the kth individual is in a particular age group, or receives a particular type of social transfer, and zero otherwise. The sum $\sum_{k=1}^{K} x_{k,j}$ therefore gives the number of individuals in the sample who are in the age group, or who receive the transfer payment.

Let the sample design weights, provided by the statistical agency responsible for data collection, be denoted s_k for $k = 1,\ldots,K$. These weights can be used to produce estimated population totals, $\hat{t}_{x|s}$ based on the sample, given by the J-element vector:

$$\hat{t}_{x|s} = \sum_{k=1}^{K} s_k x_k \tag{2}$$

Suppose that other data sources, for example census or social security administrative data, provide information about 'true' population totals, t_x. The problem is to compute new weights, w_k, for $k = 1,\ldots,K$ for which are as close as possible to the design weights, s_k, while satisfying the set of J calibration equations:

$$t_x = \sum_{k=1}^{K} w_k x_k \tag{3}$$

It is thus necessary to specify a criterion by which to judge the closeness of the two sets of weights.

In general, denote the distance between w_k and s_k as $G(w_k, s_k)$. The aggregate distance between the design and calibrated weights is thus:[5]

$$D = \sum_{k=1}^{K} G\left(w_k, s_k\right) \tag{4}$$

[4] Reference is made here to individuals, but a feature of the weights in the Household Economic Survey is that the household and individual weights are the same.
[5] Some authors, such as Folson and Singh (2000) specify the distance to be minimised as $\sum_{k=1}^{K} s_k G\left(w_k, s_k\right)$, but the present paper follows Deville and Särndal (1992).

The problem is therefore to minimise (4) subject to (3). The Lagrangean for this problem is:

$$L=\sum_{k=1}^{K} G\left(w_k, s_k\right)+\sum_{j=1}^{J} \lambda_j\left(t_{x,j}-\sum_{k=1}^{K} w_k x_{k,j}\right) \tag{5}$$

where λ_j for $j = 1,...,J$ are the Lagrange multipliers. The following subsection examines a special class of distance functions for which an iterative procedure for minimising L is developed.

A Class of Distance Functions

Consider distance functions having two features. The first is that the derivative with respect to w can be expressed as a function of w/s, and the second property is that its inverse can be obtained explicitly. Hence, $G(w_k, s_k)$ has the property:

$$\frac{\partial G\left(w_k, s_k\right)}{\partial w_k}=g\left(\frac{w_k}{s_k}\right) \tag{6}$$

The K first-order conditions for minimisation can therefore be written as:

$$g\left(\frac{w_k}{s_k}\right)=x_k' \lambda \tag{7}$$

Write the inverse function of g as g^{-1} so that if $g(w_k/s_k) = u$, say, then $w_k/s_k = g^{-1}(u)$. From(7) the k values of w are expressed as:

$$w_k=s_k g^{-1}\left(x_k' \lambda\right) \tag{8}$$

If the inverse function, g^{-1}, can be obtained explicitly, equation (8) can be used to compute the calibrated weights, given a solution for the vector, λ.

The Lagrange multipliers can be obtained by post-multiplying (8) by the vector x_k, summing over all $k = 1,...,K$ and using the calibration equations, so that:

$$t_x=\sum_{k=1}^{K} w_k x_k=\sum_{k=1}^{K} s_k g^{-1}\left(x_k' \lambda\right) x_k \tag{9}$$

Finally, subtracting $\hat{t}_{x|s}=\sum_{k=1}^{K} s_k x_k$ from both sides of (9) gives:

$$t_x-\hat{t}_{x|s}=\sum_{k=1}^{K} s_k\left\{g^{-1}\left(x_k' \lambda\right)-1\right\} x_k \tag{10}$$

The term $s_k\left\{g^{-1}\left(x_k' \lambda\right)-1\right\}$ is a scalar, and the left hand side is a known vector. In general, (10) is nonlinear in λ and so must be solved using an iterative procedure, as described in the following subsection.

An Iterative Solution Procedure

Writing $t_x-\hat{t}_{x|s} = \alpha$ the equations in (10) can be written as:

$$f_i(\lambda)=a_i-\sum_{k=1}^{K} s_k x_{k,j}\left\{g^{-1}\left(x_k' \lambda\right)-1\right\}=0 \tag{11}$$

for $i = 1,...,J$. The roots can be obtained using Newton's method. This involves the following iterative sequence, where $\lambda^{[I]}$ denotes the value of λ in the Ith iteration:[6]

$$\lambda^{[I+1]} = \lambda^{[I]} - \left[\frac{\partial f_i(\lambda)}{\partial \lambda_\ell}\right]_{\lambda^{[I]}}^{-1} \left[f(\lambda)\right]_{\lambda^{[I]}} \tag{12}$$

The Hessian matrix $\left[\partial f_i(\lambda)/\partial f_\ell\right]$ and the vector $f(\lambda)$ on the right hand side of (12) are evaluated using $\lambda^{[I]}$.

The elements are given by:

$$\frac{\partial f_i(\lambda)}{\partial \lambda_\ell} = -\sum_{k=1}^{K} s_k x_{k,i} \frac{\partial g^{-1}(x_k'\lambda)}{\partial \lambda_\ell} \tag{13}$$

which can be written as:

$$\frac{\partial f_i(\lambda)}{\partial \lambda_\ell} = -\sum_{k=1}^{K} s_k x_{k,i} x_{k,\ell} \frac{dg^{-1}(x_k'\lambda)}{d(x_k'\lambda)} \tag{14}$$

Starting from arbitrary initial values, the matrix equation in (12) is used repeatedly to adjust the values until convergence is reached, where possible.

As mentioned earlier, the application of the approach requires that it is limited to distance functions for which the form of the inverse function $g^{-1}(u)$, can be obtained explicitly, given the specification for $G(w,s)$. Hence, the Hessian can easily be evaluated at each step using an explicit expression for $dg_k^{-1}(x_k'\lambda)/d(x_k'\lambda)$. As these expressions avoid the need for the numerical evaluation of $g^{-1}(x_k'\lambda)$ and $dg_k^{-1}(x_k'\lambda)/d(x_k'\lambda)$ for each individual at each step, the calculation of the new weights can be expected to be relatively quick, even for large samples.[7] However, a solution does not necessarily exist, depending on the distance function used and the adjustment required to the vector $t_x - \hat{t}_{x|s}$.

Some Distance Functions

Consider the chi-squared distance measure, where the aggregate distance is given by:

$$G(w_k, s_k) = \frac{1}{2}\sum_{k=1}^{K} \frac{(w_k - s_k)^2}{s_k} \tag{15}$$

Here, $g(w_k/s_k) = w_k/s_k - 1$, and it can be shown that an explicit solution exists with:

$$w_k = s_k\left(1 + x_k'\lambda\right) \tag{16}$$

for $k = 1,...,K$ and:

$$\lambda = \left[\sum_{k=1}^{K} s_k x_k x_k'\right]^{-1} \left(t_x - \hat{t}_{x|s}\right) \tag{17}$$

---

[6] The approach described here differs somewhat from other routines described in the literature, for example in Singh and Mohl (1996) and Vanderhoeft (2001). However, it provides extremely rapid convergence.

[7] Using numerical methods to solve for each $g^{-1}(u)$ and $dg^{-1}(u)/du$, for $u = x_k'\lambda$, for every individual in each iteration, would increase the computational burden substantially.

where the term in square brackets on the right hand side of (17) is a J by J square matrix.[8]

One reason why the chi-squared distance function produces a solution is that no constraints are placed on the size of the adjustment to each of the survey weights. It is therefore also possible for the calibrated weights to become negative. However, Deville and Särndal (1992) suggested the following simple modification to the chi-squared function, although the explicit solution for the chi-squared case is no longer available and the iterative method must be used.

Suppose it is required to constrain the proportionate changes to certain limits, different for increases compared with decreases in the weights. Define r_L and r_U such that $r_L<1<r_U$. The objective is to ensure that, for increases, the proportionate change, $w/s - 1$, is less than $r_U - 1$ or that $r_U > w/s$. For decreases, the aim is to ensure that $1 - w/s$ (or the negative of the proportional change) is less than $1 - r_L$ so that $r_L < w/s$.

For the chi-squared distance function, $g^{-1}(u) = 1 + u$, where $u = x'\lambda$ and $g^{-1}(u)$ solves for w/s. Hence if $g^{-1}(u) = w/s$ is outside the specified range, it is necessary to set it to the relevant limit, either r_U or r_L, rather than allow it to take the value generated. Since $g^{-1}(u) - 1 = w/s - 1 = u$ the limits are exceeded if $u < r_L - 1$ and if $u > r_U - 1$. In each case where the value of $g^{-1}(u)$ has to be set to the relevant limit, the corresponding value of $dg^{-1}(u)/du$ is zero. This approach ensures that weights are kept within the range, $r_L s_k < w_k < r_U s_k$. Hence, negative values w of are avoided simply by setting r_L to be positive.[9]

It is not necessary to start from a specification of $G(w,s)$, since the solution procedure requires only an explicit form for the inverse function $g^{-1}(u)$, from which its derivative can be obtained. Deville and Särndal (1992) suggested the use of an inverse function $g^{-1}(u)$ of the form:[10]

$$g^{-1}(u) = \frac{r_L(r_U - 1) + r_U(1 - r_L)\exp \alpha u}{(r_U - 1) + (1 - r_L)\exp \alpha u} \tag{18}$$

where r_L and r_U are as defined above and:

$$\alpha = \frac{r_U - r_L}{(1 - r_L)(r_U - 1)} \tag{19}$$

---

[8] Write (16) as $w_k = s_k\left(1 + \lambda' x_k\right)$ and (17) as $\lambda' = \left(t_x - \hat{t}_{x|s}\right)' T^{-1}$ with as the symmetric matrix $\sum_{k=1}^{K} s_k x_k x_k'$. Given sample observations on the variable y_k, an estimate of the population total $\hat{t}_y$, can be obtained as $\sum_{k=1}^{K} w_k y_k$. Substituting for w_k gives the result in Deville and Särndal (1992, p.377) that $\hat{t}_y = \sum_{k=1}^{K} s_k y_k + \left(t_x - \hat{t}_{x|s}\right)' B$, where $B = T^{-1} \sum_{k=1}^{K} s_k x_k y_k$. This provides the link between reweighting and the Generalised Regression (GREG) estimator. The production of asymptotic standard errors is often based on this estimator, in view of the result that other distance functions are asymptotically equivalent; see Deville and Särndal (1992, p.378).

[9] This is much more convenient than imposing inequality constraints and applying the more complex Kuhn-Tucker conditions. Also, it is desirable to restrict the extent of proportional changes even where they produce positive weights.

[10] Singh and Mohl (1996), in reviewing alternative calibration estimators, refer to this 'inverse logit-type transformation' as a Generalised Modified Discrimination Information method.

Thus $g^{-1}(-\infty) = r_L$ and $g^{-1}(\infty) = r_U$, so that the limits of w/s are r_L and r_U. This function therefore has the property that adjustments to the weights are kept within the range, $r_L s_k < w_k < r_U s_k$, although, unlike the chi-squared modification, no checks have to be made during computation.

The derivative required in the computation of the Hessian is therefore:

$$\frac{dg^{-1}(u)}{du} = g^{-1}(u)\left\{r_U - g^{-1}(u)\right\} \frac{(1-r_L)\alpha \exp \alpha u}{(r_U - 1) + (1-r_L)\exp \alpha u} \tag{20}$$

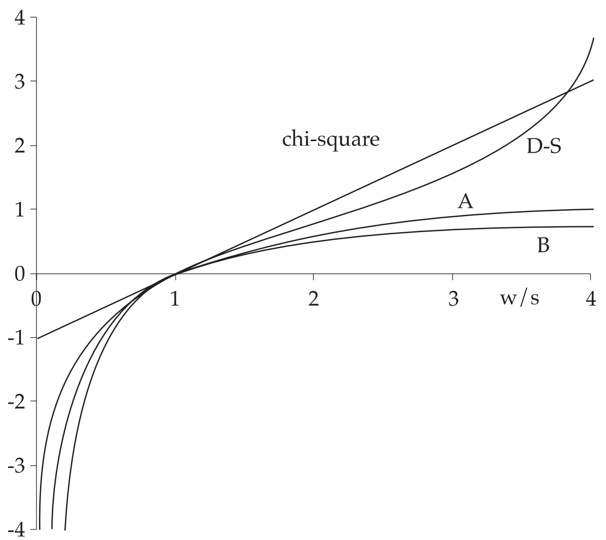
Since $g^{-1}(u)$ solves w/s for (18) can be rearranged, by collecting terms in $\exp \alpha u$, to give:

$$\frac{\frac{w}{s} - r_L}{1 - r_L} = \frac{r_U - \frac{w}{s}}{r_U - 1} \exp \alpha u \tag{21}$$

so that the gradient of the distance function is:

$$g\left(\frac{w}{s}\right) = u = \frac{1}{\alpha}\left[\log\left(\frac{\frac{w}{s} - r_L}{1 - r_L}\right) - \log\left(\frac{r_U - \frac{w}{s}}{r_U - 1}\right)\right] \tag{22}$$

Figure 1 Alternative Gradient Functions



The special nature of this gradient function is illustrated by the line D-S in Figure 1, which shows the profile of (22) for $r_U = 4.1$ and $r_U = 0.01$. The restriction of w/s to the range specified is evident.[11] Figure 1 also shows the function $g(w/s)$ for two other cases mentioned by Deville and Särndal (1992). Case A uses $g^{-1}(u) = \left(1 - \frac{u}{2}\right)^{-2}$, and case B has $g^{-1}(u) = (1-u)^{-1}$.[12] In all cases,

[11] The distance function itself is given by integrating (22) with respect to w. giving s/α multiplied by $G(w,s) = \left(r_U - \frac{w}{s}\right)\log\left(\frac{r_U - \frac{w}{s}}{r_U - 1}\right) + \left(\frac{w}{s} - r_L\right)\log\left(\frac{\frac{w}{s} - r_L}{1 - r_L}\right)$ plus a term $(r_U - r_L)s/\alpha$, which, since it is a constant, may be dropped without loss. The derivation is given in Creedy (2003).

the slope is zero, corresponding to a turning point of the distance function, when $w/s = 1$. Given the quadratic U-shaped nature of the chi-squared distance function, the gradient increases at a constant rate, being negative in the range $w/s < 1$. Cases A and B also imply U-shaped distance functions, but with the gradient increasing more sharply for $w/s < 1$ and more slowly than the chi-square function in the range $w/s > 1$.

3. TaxMod Estimates

This section compares, for each area of expenditure, the estimates obtained using the New Zealand Treasury microsimulation model, TaxMod, with unpublished data on the 'actual' expenditures. The latter data were obtained from the Inland Revenue Department and the Ministry of Social Development. However, they are obtained from samples taken from the basic beneficiary data, in view of the difficulty of obtaining complete information at the level of aggregation required. The most recent Household Economic Survey (HES) data are for the 2000-01 year.

One role of a microsimulation model is to examine, along with aggregate cost estimates, the extent to which particular groups in the population are likely to gain or lose from a tax reform. For this reason it is important to ensure that the model provides a good representation of the extent to which expenditures on different types of benefit go to different types of family. Table 1 summarises benefit expenditures for 2000-01, disaggregated into a variety of household types. The values reported for TaxMod use the weights provided by Statistics New Zealand.[13]

The final column of table 1 shows the percentage difference between actual values and TaxMod estimates, calculated as 100x(actual-TaxMod)/TaxMod: hence negative values indicate an overstatement by TaxMod. The table shows that TaxMod overestimates aggregate expenditure on the Unemployment Benefit by 4.1 per cent, underestimates expenditure on the Domestic Purposes Benefit by 1.9 per cent, and underestimates aggregate expenditure on the Invalids' and Sickness Benefits by 38.6 and 18 per cent respectively. For all benefit categories, TaxMod tends to underestimate expenditure on single income families and, in contrast, overestimates expenditure on partnered families. In some cases, particularly Domestic Purposes Benefit recipients without children, this general pattern did not apply, possibly reflecting the small size of the sample for certain demographic groups or the difficulty of modelling certain population characteristics. TaxMod underestimates total expenditure on (combined) Family Support, Child and Family Tax Credits to single families and, in contrast, overestimates expenditure on Family Support and the Child Tax Credit to partnered families.

TaxMod computes benefit expenditures on the assumption that all those who are eligible actually claim their full entitlement. However, it is known that benefit take-up rates are often less than 100 per cent. This feature would produce a consistent upward bias in TaxMod estimates, which is not evident

[12] Deville and Särndal (1992) discuss the use of a normalisation whereby g^{-1} is set to some specified value, but this is not necessary for the approach.
[13] These are integrated weights, not the original weights. For a discussion of the use of integrated weighting, as described by Lemaître and Dufour (1987), by Statistics New Zealand, see StatsNZ (2001).

here.[14] The differences were thus judged sufficiently large to warrant reweighting.

Table 1 Benefit Expenditure by Family Types (2000-01)

		TaxMoc	*I Share*	*Actual*	*Share*	*%Diff*
		($m)	*(%)*	*($m)*	*(%)*	
Unemployment Benefit						
Single	no children	648	48.7	814	63.7	25.6
	1+ children	58	4.4	80	6.3	37.9
Couple	No children	175	13.1	139	10.9	-20.6
	1 child	194	14.6	77	6.0	-60.3
	2 children	148	11.1	79	6.2	-46.6
	3+ children	108	8.1	88	6.9	-18.5
All		1,331	100.0	1277	100.0	-4.1
Domestic Purposes Benefit						
Single	No children	120	9.8	43	3.4	-64.2
	1 child	465	38.0	550	44.1	18.3
	2 children	388	31.7	404	32.4	4.1
	3+children	250	20.4	250	20.0	0
Others		1	0.1	0	0	0
All		1,224	100.0	1,247	99.9	1.9
Invalid Benefit						
Benefit						
Single	No children	287	62.5	423	66.5	47.4
Couple	No children	81	17.6	115	18.1	42.0
Others		91	19.8	98	15.4	7.7
All		459	99.9	636	100.0	38.6
Sickness Benefit						
Benefit						
Single	No children	138	51.7	195	61.9	41.3
Couple	1+ children	58	21.7	55	17.5	-5.2
Others		71	26.6	65	20.6	-8.5
All		267	100.0	315	100.0	18.0
Family *i* Support, and Family Tax Credits						
Single	1 child	178	15.5	203	20.1	14.0
	2 children	176	15.3	217	21.5	23.3
	3+ children	175	15.2	184	18.3	5.1
Couple	1 child	88	7.6	60	6.0	-31.8
	2 children	216	18.8	128	12.7	-40.7
	3+ children	315	27.5	216	21.4	-31.4
All		1,174	100.0	1,008	100.0	-14.1

4. Re-Weighted Estimates

The previous section has shown that, for some of the household types, the discrepancy between TaxMod estimates and actual expenditure is substantial. This suggests that in reweighting the Household Expenditure Survey, it is important to use calibration values relating to these particular types.[15] The calibration requirements used for reweighting are presented in subsection 1. The revised weights are discussed in subsection 2.

[14] It would not be appropriate to adjust the sample weights if it were felt that the main problem related to imperfect take-up of benefits.
[15] Nascimento Silva and Skinner (1997) examined variable selection in general, but in the present context the variables naturally arise.

Calibration Conditions

Tables 2, 3 and 4 show the calibration conditions used in reweighting: the required population totals are given in the second column of each table, under the heading 'required total'. These cover respectively the numbers in each family type, the number of benefit recipients in each demographic group, and the number of individuals in each age group. To avoid singularities, it was of course necessary to omit one category from each of the classes. The tables show only those calibration conditions actually used. The numbers produced by TaxMod, using the weights provided by Statistics New Zealand, are shown in the third column of each table. The differences between the required and estimated totals, shown in the final column of each table, are substantial. These reflect a larger population size combined with population ageing, an increase in the number of singles, and particularly singles receiving Domestic Purposes Benefit, Unemployment Benefit and Invalid's Benefit.

Table 2 Calibration: Family Composition

Demographic Group	*Required Total*	*Stats NZ Weights*	*Difference*
Couples			
1 child	343258	338025	5233
2 children	537178	614249	-77071
3 children	279735	338800	-59065
4 children	98436	123542	-25107
5+children	55267	37108	18159
Single Persons			
no children	1133969	650062	483907
1 child	142875	111423	31452
2 children	140624	92212	48412
3 children	73284	48016	25268
4 children	31389	17002	14387
5+ children	19508	13300	6207

Table 3 Calibration: Number of Benefit Recipients

Unemployment Benefit	*Required Total*	*Stats NZ Weights*	*Difference*
Single person	104292	58735	45557
Sole parent 1child	6400	2273	4127
Couple no child	11045	19628	-8583
Couple one child	4327	12349	-8022
Couple 2+children	9206	19560	-10354
Domestic Purposes Benefit			
No children	10285	6891	3394
One child	52988	32231	20757
Two+children	57647	36140	21507
Invalidity Benefit			
Single	61343	20345	40998
Couple	13186	24866	-11680
Sickness Benefit			
Single	42137	14858	27279
Couple	9908	6977	2931
widow's beneficiaries			
All	9870	8026	1844

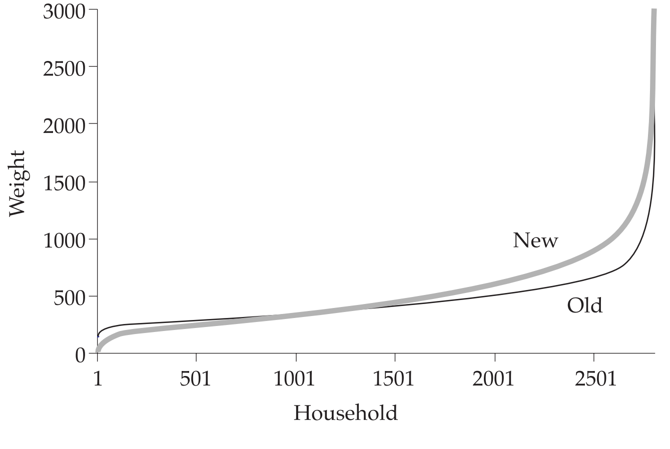
Table 4 Calibration: Age Distribution

	Required Total	*Stats NZ Weights*	*Difference*
Males			
5-9	145204	170507	-25303
10-14	150403	134372	16031
15-19	137214	87803	49411
20-24	116565	81554	35011
25-44	516856	467126	49730
45-59	353453	296280	57173
60-74	199651	194141	5510
Females			
0-4	127864	132633	-4769
5-9	138368	112789	25578
10-14	142813	122452	20361
15-19	133253	94175	39078
20-24	116926	85468	31458
25-44	568121	543435	24686
45-59	367692	313687	54005
60-74	211984	190247	21737
Males and Females			
75+	192415	153097	39318

Revised Weights

The variation in the survey weights provided by Statistics New Zealand for the period 2000/01 is illustrated by the solid line in figure 2, where the weights are arranged in ascending order for the Household Economic Survey sample of 2808 households. The number on the horizontal axis thus refers to the rank of the household. It can be seen that the majority of these weights are within a fairly narrow range, although some are substantially higher, suggesting a considerable degree of under-representation of these household types in the sample.

Figure 2 Statistics New Zealand and New Weights



The iterative reweighting method described earlier was applied using the various distance functions described. However, it was found that no solution exists for the Deville and Särndal (1992) function, whatever limits are imposed on the proportional changes in weights. The procedure produced solutions for the modified chi-squared case, with the upper and lower ratios set to 6 and 0.06 respectively. The approach used was to start with broad limits and 'work inwards' so long as solutions are available. The iterative method quickly reveals when a solution is not possible. As mentioned earlier, convergence using the Newton method is extremely rapid. Figure 2 also shows, as the dashed line, the new weights, also arranged in ascending order. Compared with the initial weights, the increase in the population size is evident, with most of the weights increasing.

The range of upper and lower limits imposed on the ratios of old to new weights appears to be very large. However, it was found that very few of the new weights actually reach those limits. This can be seen from Figure 3 and particularly Figure 4, which show the revised weights and the ratio of new to old weights, with the households ranked in the same order as in the solid line in Figure 2.

Figure 3 Revised Weights

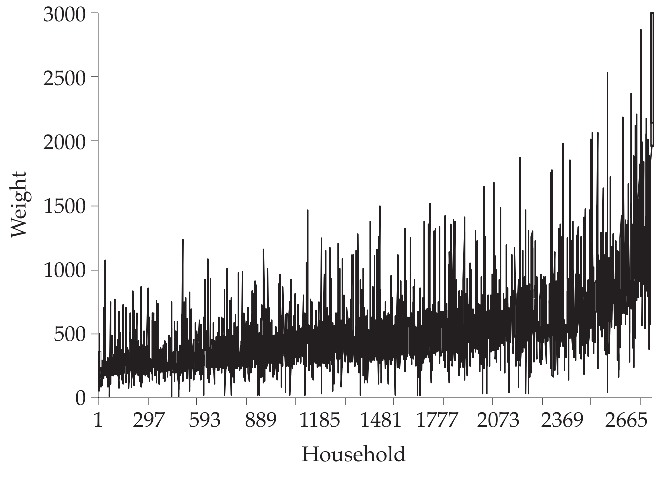


Figure 4 Ratio of Revised to Initial Weights

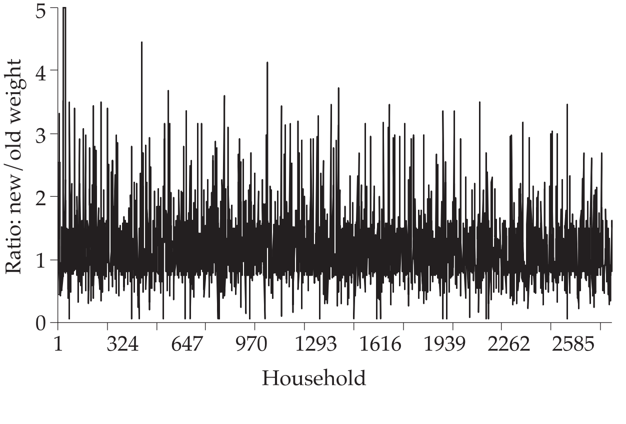


Table 5 Re-Weighted Benefit Expenditures by Family Types

		TaxMoc ($m)	Share (%)	Actual ($m)	Share (%)	%Diff
Unemployment Benefit						
Single	No children	815	63.4	814	63.7	-0.1
	1+ children	80	6.2	80	6.3	0.0
Couple	No children	138	10.7	139	10.9	0.7
	1 child	86	6.7	77	6.0	-10.5
	2 children	166	12.9	167	13.1	0.6
All		1,286	100	1,277	100.0	-0.7
Domestic Purposes Benefit						
Single	No children	44	3.5	43	3.4	-2.3
	1 child	550	44.1	550	44.1	0.0
	2 children	391	31.3	404	32.4	3.3
	3+ children	263	21.1	250	20.0	-4.9
All		1,249	100	1,247	99.9	-0.2
Invalid Benefit						
Single		463	72.7	463	72.7	0.0
Couple		174	27.3	174	27.3	0.0
All		637	100.0	636	100.0	-0.2
Sickness Benefit						
Single		216	68.8	216	68.6	0.0
Couple		98	31.2	99	31.4	1.0
All		314	100.0	315	100.0	0.3
Family Support, Child and Family Tax Credit						
Single	1 child	185	16.6	203	20.1	9.7
	2 children	181	16.2	217	21.5	19.9
	3+ children	183	16.4	184	18.3	0.5
Couple	1 child	69	6.2	60	6.0	-13.0
	2 children	197	17.6	128	12.7	-35.0
	3+ children	302	27.0	216	21.4	-28.5
All		1,116	100.0	1,008	100.0	-9.7

The calibrations are based on numbers of individuals and households falling into the various categories, rather than total expenditures. It is therefore not obvious that aggregate expenditure levels are significantly improved. The implications of using the revised weights for estimated expenditures in each demographic group are reported in table 5. It can be seen that in most cases the TaxMod estimates are much closer to the 'actual' values. However, some concern remains over Family Support, Child and Family Tax Credits.

5. Income Distributions

Reference has been made briefly to the important concern regarding the possible effects of reweighting on important variables which are not part of the calibration exercise.[16] This section examines the income distribution before and after reweighting.

Figure 5 compares the distributions of annual gross income obtained using the two sets of weights.[17] In view of the calibration conditions, it is not

[16] This point was stressed by, for example, Klevmarken (1998).
[17] For present purposes each income has been rounded to the nearest multiple of $2000 and the distribution is truncated at $150,000.

surprising that the reweighted distribution has more people with benefit-level incomes than with the original weights. The compensating reduction in frequencies is spread over quite a wide range of higher incomes. An effect of the chosen reweighting is to increase the total number of people in the population: this is of course not shown in the figure.

Figure 5 Frequency Distribution of Income

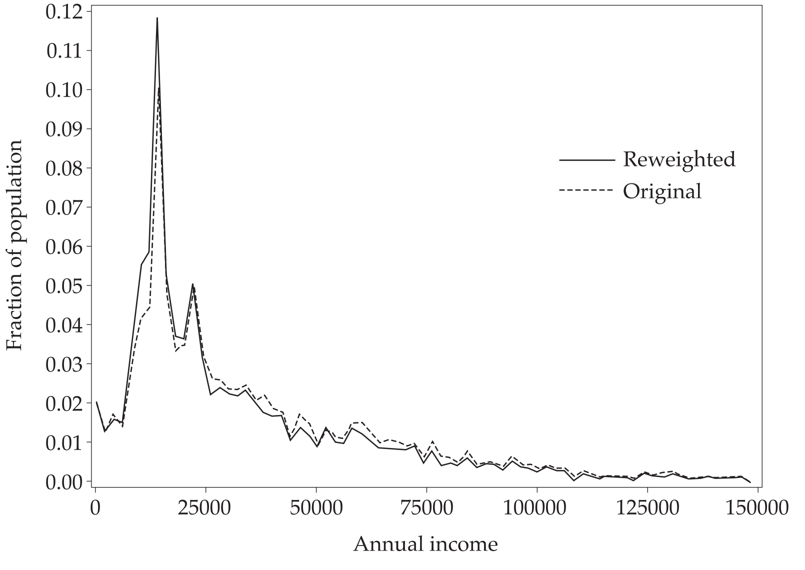


Figure 6 Cumulative Income Distribution

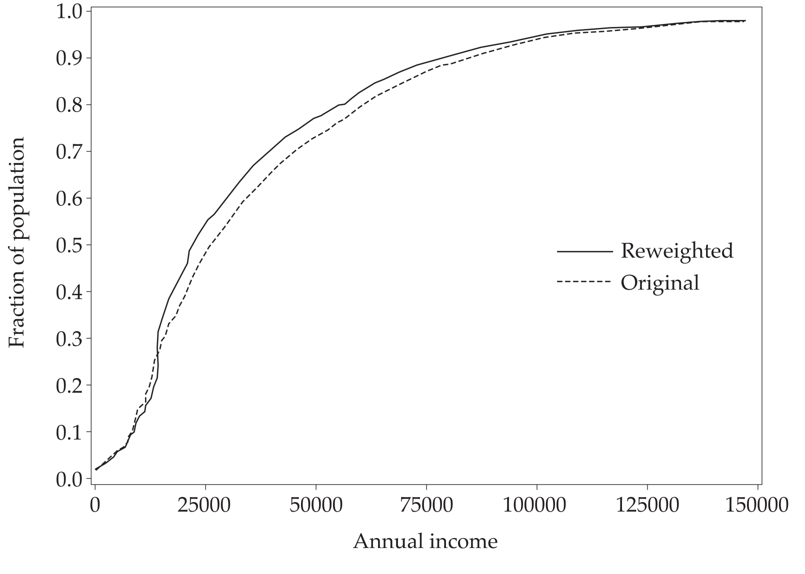


Figure 7 Differences in Proportions

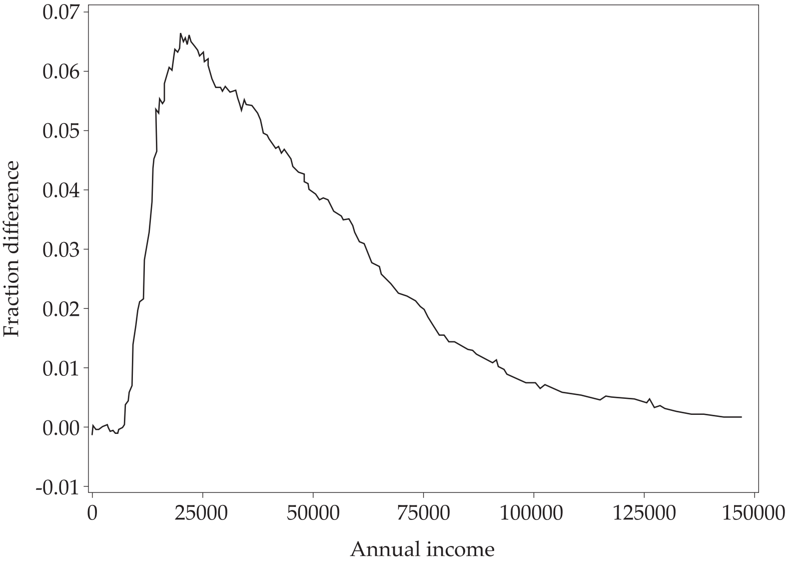


Figure 6 compares the cumulative income distributions before and after reweighting. The fact that the reweighted income distribution is weighted more heavily towards low incomes than in the original data is also revealed in this figure. It may not be obvious just how wide the gap can be between the two curves, so figure 7 shows the vertical differences at each income level. The vertical scale measures the percentage of the total population; that is, the peak of 6.5 per cent does not mean that the reweighted numbers are 6.5 per cent greater at an annual income of $20,000, but rather that the reweighted figures have an additional 6.5 per cent of the total population earning $20,000 or less, compared with the original figures. In view of the calibration conditions used, these changes in the income distribution appear to be quite reasonable.

6. A Policy Simulation

Having obtained new weights, it is useful to consider their effects on a policy simulation. For present purposes it is best to specify a very simple policy change, the effects of which are transparent. Suppose the New Zealand income tax rates of 33 and 39 per cent are raised to 35 and 41 per cent respectively. At the same time, family support rates are increased by $10 per week. Clearly, all income tax payers who are not in receipt of benefits will lose from this reform. Summary information about the reform, using both the Statistics New Zealand weights and the revised weights, is given in table 6.

The first block of the table decomposes the winners and losers by family type. Using the revised weights, the policy produces more winners who are single parents with two or more children; there are 71 thousand who gain using the new weights compared with 51 thousand families under the initial weights. However, there are fewer couples with two or more children who gain (111 compared with 126 thousand). These differences are also revealed when considering the net changes in government expenditure, shown in the second block of the table for the same household types. Government expenditure on single parents with two or more children increases by more, while that on couples with two or more children increases by less, when the revised weights are compared with the old. In total, with the new weights, the policy change raises less net revenue (or has a lower reduction in net costs) compared with the old weights: the net revenue change is $14.6m compared with $38.2m. This change is clearly consistent with the increase in the number of beneficiaries reflected in the revised weights.

Table 6 A Policy Reform Using Alternative Weights

	Stats NZ Weights				*Revised Weights*			
	Gain	*NC*	*Loss*	*All*	*Gain*	*NC*	*Loss*	*All*
	Families Affected (000s)							
Single no child	0	699	112	810	0	866	132	998
Single 1 child	55	1	6	61	67	0	2	68
Single 2+child	51	0	1	52	71	0	0	71
Couple no child	0	295	175	470	0	282	160	442
Couple 1 child	34	27	57	118	31	23	53	107
Couple 2+child	126	13	120	259	111	13	104	228
Overall	266	1034	471	11	280	1184	450	1914
	Changes in Costs ($M)							
Single no child	0	0	-52.7	-52.7	0	0	-64.2	-64.2
Single 1 child	27.7	0	-2.0	25.7	33.6	0	-0.6	33.0
Single 2+child	64.4	0	-0.9	63.6	91.0	0	-0.1	90.9
Couple no child	0	0	-121.1	-121.1	0	0	-112.2	-112.2
Couple 1 child	16.5	0	-34.6	-18.1	15.4	0	-34.3	-18.9
Couple 2+child	167.52	0	-103.2	64.3	150.5	0	-93.7	56.81
Overall	276.2	0	-314.4	-38.2	290.5	0	-305.1	-14.6
	Families Affected (000s)							
Unemployment	20	84	1	105	24	136	1	161
DPB	69	9	0	77	116	3	0	119
Invalids Benefit	7	28	0	35	7	53	0	60
Sickness Benefit	4	15	0	20	8	30	0	38
Widows Benefit	2	7	0	9	1	9	0	10
NZ Super	1	309	24	334	1	328	24	353
None of above	162	583	446	2122	122	624	425	1172
Overall	266	1034	471	11	280	1184	450	1914
	Changes in Costs ($M)							
Unemployment	22.1	0	-0.1	22	27.9	0	-0.2	27.7
DPB	63.6	0	0	63.6	109.2	0	0	109.2
Invalids Benefit	6.1	0	0	6.1	6.1	0	0	6.1
Sickness Benefit	3.7	0	0	3.7	6.3	0	0	6.3
Widows Benefit	1.2	0	0	1.2	0.6	0	0	0.6
NZ Super	2.3	0	-15.9	-13.6	3.5	0	-15.6	-12.0
None of above	177.1	0	-298.4	-121.3	136.8	0	-289.4	-152.6
Overall	276.2	0	-314.4	-38.2	290.5	0	-305.1	-14.6

The changes are decomposed by benefit type in the last two blocks of the table. There are many more families who gain from the reform and are in receipt of the Domestic Purposes Benefit, when the new weights are used

(116 compared with 69 thousand families). This translates into a larger increase in the cost of the DPB of $109.2m with the new weights compared with $63.6m under Statistics New Zealand weights. The increase in revenue arising from the large number of losers in the 'none of the above' categories comes from those who are taxpayers only. It is clear that judgements about the likely effects of the policy are influenced by the weights used.

7. Conclusions

This paper has reported a reweighting exercise for the New Zealand Household Economic Survey, which is the basis of the Treasury's microsimulation model, TaxMod. Comparisons of benefit expenditures in a variety of demographic groups, along with population data, showed that TaxMod estimates often differed substantially from estimated totals based on administrative data, when the weights provided by Statistics New Zealand were used. After describing the basic method used to compute new weights, the calibration requirements were reported. These relate to the age structure of the population and the number of beneficiaries for Unemployment Benefits, Domestic Purposes Benefit, Invalid's and Sickness benefits and Family Support and Tax Credits. The revised weights and expenditure estimates were reported and the resulting distribution of income was examined. The new weights were found to produce much improved estimates of total expenditure on the various categories within a range of demographic groups, without distorting the resulting income distribution. Finally a policy simulation, involving an increase in income tax rates and family support payments, demonstrated that reweighting can indeed have an important effect on estimates of the implications of a planned tax reform.

References

Andersson, C. and Nordberg, L. (1998), A User's Guide to CLAN 97, *Statistics Sweden.*

Bell, P. (2000), Weighting and standard error estimation for ABS household surveys. *Paper prepared for ABS Methodology Advisory Committee: Australian Bureau of Statistics.*

Creedy, J. (2003), Survey Reweighting for Tax Microsimulation Modelling. *New Zealand Treasury Working Paper*, No. 03/17.

Deville, J.-F. and Särndal, C.-E. (1992) Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 87, 376-382.

Folson,R.E.Jnr. and Singh, A.C. (2000), The Generalized Exponential Model or Sampling Weight Calibration for Extreme Values, Non-response and Post-stratification. *Proceedings of the Survey Research Methods Section: American Statistical Association*. http://www.amstat.org/sections/srms/proceedings/papers/2000_099.pdf.

Klevmarken, N.A. (1998), Statistical Inference in Micro-simulation Models: Incorporating External Information, *Uppsala University Department of Economics Working Paper,* http://www.nek.uu.se/Pdf/1998wp20.pdf.

Lemaître, G. and Dufour, J. (1987), An Integrated Method for Weighting Persons and Families, *Survey Methodology*, 13, 199-207.

Nascimento Silva, P.L.D. and Skinner, C. (1997), Variable Selection for Regression Estimation in Finite Populations, *Survey Methodology*, 23.

Singh, A.C. and Mohl, C.A. (1996), Understanding Calibration Estimators in Survey Sampling, *Survey Methodology*, 22, 107-115.

Statistics New Zealand (2001), Information Paper: The Introduction of Integrated Weightingto the 2000/2001 Household Economic Survey, *Statistics New Zealand*

Vanderhoeft,C.(2001), Generalised Calibrationat Statistics Belgium, *Statistics Belgium Working Paper,* No. 3.