

Computation of the Compensating Variation within a Random Utility Model Using GAUSS Software

Marilena Locatelli, Steinar Strøm

University of Turin and Child Centre, Turin, Italy,

Frisch Centre, Oslo, Norway

E-mail: marilena.locatelli@unito.it, steinar.strom@econ.uio.no

Received February 23, 2011; revised April 10, 2011; accepted April 22, 2011

Abstract

To evaluate a tax reform in terms of change in household welfare one possibility is to estimate the compensating variation using a suitable model to assess the change in the household utility. When a random utility model is used, the computation of compensating variation is not straightforward, particularly when utility is not linear in household income. It can be carried out using a methodology recently proposed in the literature. In this paper we describe a software instrument, implemented using GAUSS programming language for computing the compensating variation to evaluate the 1991 tax reform introduced in Norway. The program is flexible and adaptable to different tax systems and different reference years.

JEL Classification: C63, B21

Keywords: Compensating Variation, Gauss Application, Tax System Evaluation

1. Introduction

In 1992 the Norwegian tax system was reformed towards lower and less progressive tax rates, with a reduction in the total tax revenue. In the next following years the tax structure was kept nearly unchanged. To evaluate a tax reform in terms of change in household welfare one possibility is to estimate the compensating variation (CV) using a suitable model to assess the utility of the households.

The theoretical framework and the empirical results of the evaluation of the above tax reform are reported in [1] in which there is also an extensive literature review on this subject. Novelty of the paper is the use of a random utility labor supply model, taking into accounting also sectoral choices (public and private), to assess the impact of tax reforms on household welfare. Some partial details on the algorithm implementation were first given in [2].

The computation of CV is not straightforward in a random utility model, in particular when utility is not linear in household income. A random utility function implies that the expenditure function is also random. Until recently, no analytic formulas have been available for calculating the distribution of CV. However, [3] have developed analytic formulas for this purpose, and we apply their methodology to calculate the distribution of CV and the mean and variance of this distribution.

What we thus do is to calculate the expected value of CV for each household and its distribution in the population.

The purpose of this note is to describe the software instrument, implemented using Gauss software package [4], to evaluate the above tax reform in terms of change in household welfare within a random utility model. The program is flexible and adaptable to different tax systems and different reference years.

The paper is organized as follows. In the next Section, the data used in estimations are described. The model is explained in Section 3. In Section 4 we analyze the main steps of the procedure and in Section 5 the program structure is outlined. The results are reported in Section 6. Section 7 draws some conclusions.

The Table upon the estimation is reported in Appendix A. The interpolation method followed in the computation of CV is described in Appendix B. In Appendix C we report the GAUSS program flow. The complete program code can be downloaded [2].

2. The Data

Data on the labor supply of married women in Norway used in this note consist of a merged sample from “Survey of Income and Wealth, 1994”, Statistics Norway

(1994) and “Level of Living Conditions, 1995”, Statistics Norway (1995). Data cover married couples as well as cohabiting couples with common children. The age of the spouses ranges from 25 to 64. None of the spouses are self-employed and none of them are on disability or other type of benefits. All taxes paid are observed and in the assessment of disposable income, all details of the tax system are accounted for.

The size of the sample used in estimating the labor supply model is 810 (referred to as *n_record* in the program).

3. The Model

To evaluate the tax reform of 1992 we calculate the change in household welfare. One way is to apply the measure of CV. The calculation of CV is not straightforward in a random utility model when utility is not linear in household income. A random utility function implies that the expenditure function is also random. A general treatment of this issue was undertaken by [3] while in [1] the method was adapted to the change in household welfare with labor supply random utility model.

What we do is to calculate the expected value of CV, $E[CV]$, for each individual and thereafter the distribution of this value in the population from which we can derive mean, median etc.

We will assume that the utility function has the structure

$$U(C, h, z) = v(C, h)\varepsilon(z), \quad \text{for } z = 0, 1, 2, 3, \dots \quad (1)$$

where C is the household disposable income, which equals the sum of the after-tax labor income of husband and wife plus the after tax capital income and public transfer like child allowances; h are hours of work. In details: 0 (not working), 315, 780, 1040, 1560, 1976, 2340, 2600 (both for public and private sector); $v(\cdot)$ is a deterministic function and $\varepsilon(z)$ is a positive random taste shifter. The taste shifter accounts for unobserved individual characteristics and unobserved job-specific attributes z . $\{\varepsilon(z)\}$, are independently distributed with c.d.f. $\exp(-x^{-1}), x > 0$.

From Equation (1) we get the following implicit definition of the CV when the tax regime of 1991 (prior to the tax reform) is compared to the tax regime of 1994 (after the tax reform):

$$\begin{aligned} U(C, h, \varepsilon(z)|\text{Tax regime 1991}) = \\ U(C - CV, h, \varepsilon(z)|\text{Tax regime 1994}) \end{aligned} \quad (2)$$

In Equation (1) we have suppressed the subscript of the individual and we should also keep in mind that the choice of each individual is to choose to work or not, and given work, to choose sector and hours of work, given

the job opportunity sets and the budget constraints under the different tax regimes and CV. In the calculation of the expected value of CV we take this choice structure into account. If an individual benefits from the tax reform, the expected value of CV for this individual is positive, meaning that this amount has to be subtracted from household income under the 1994 tax regime in order to make the individual indifferent between the two tax regimes.

To proceed with the calculation we need some notation. Note first that the deterministic part of the utility function can be written $v_i = v(C(h_i), h_i|X)$ (v_i referred to as consumption function in the following), where i denotes the different job alternatives. In the examined case the choice alternatives (*na*) are 15: $i = 1$ the individual is not working, and $i = 2, 3, \dots, 8$ denote hours of job in the public sector, while $i = 9, \dots, 15$ denote hours of job in the private sector. X is a vector of all exogenous characteristics.

Now let

$$v_i^0 = v(C(h_i|\text{Tax regime 1991}), h_i|X) \quad (3)$$

and let

$$v_i^*(y) \equiv v(C^*(h_i, y), h_i|X) \quad (4)$$

where $C^*(h_i, y) = F(h_i) + y$, $F(h_i) = w_i h_i - T(w_i h_i)$ and $T(\cdot)$ is the tax function for 1994, w_i is the wage rate.

Let $E[CV]$ be the expected value of the compensating variation, which can be calculated for each individual as follows [3]:

$$E[CV] = I^* - \sum_i^{15} \left[v_i^0 g_i b_i \int_0^{\tilde{y}_i} \frac{dy}{\sum_{i=1}^{15} \max[v_i^0 g_i b_i, v_i^*(y) g_i b_i]} \right] \quad (5)$$

where, according to estimates given in [5] and reported in Appendix A,

$$\begin{aligned} g_i &= 1 \text{ for all } i \text{ except for } i = 4, 6, 11, 13 \\ g_4 &= \exp(0.68) \\ g_6 &= \exp(1.58) \\ g_{11} &= \exp(0.80) \\ g_{13} &= \exp(1.06) \\ b_1 &= 1 \\ b_i &= \exp(-4.20 + 0.22X_4); \text{ for } i = 2, 3, \dots, 8 \\ b_i &= \exp(1.14 - 0.33X_4); \text{ for } i = 9, 10, \dots, 15. \end{aligned}$$

and \tilde{y}_i is given by the following equation:

$$v_i^0 = v_i^*(\tilde{y}_i) \quad (6)$$

I^* (see Equation (5)) equals the sum of the after tax income of husband earnings and capital income, plus child allowances. The tax reform of 1992 is a combination of a change of the tax structure and reduction in tax revenues.

4. Steps of the Procedure

In 1992 the Norwegian tax system was reformed towards lower and less progressive tax rates with a reduction in the total tax revenue. We have thus organized the program to allow also the computation of the expected E[CV] between the 1994 tax system and the flat tax system (in our case 29%) that equals tax revenue of year 1994.

4.1. The Algorithm

To calculate E[CV] the following steps are required:

1) Load the matrix file, previously prepared¹, with the values v_i^0 (the household's deterministic utility (i.e. consumption function)) for the reference year 1991 and working hours h_i .

$$v_i^0 = v(C(h_i | Tax \ function \ 1991), h_i | X) \quad (7)$$

where v_i^0 is a $n_record \times na$ matrix, being n_record the number of records of the matrix file, and na the number of alternatives of the choice set (i.e. 15 alternatives).

We also calculate E[CV] with reference to a flat tax system (in our case the revenue neutral tax rate simulated on the choice model is 29%). The model under the flat tax system gives the reference values when the 1994 tax regime is evaluated against the flat tax system.

Now, load the matrix file with the averaged values v_i^0 (the household deterministic utility) under the flat tax system.

2) Load the data sets including:

- a) the variables (disposable income, etc.) for the tax system 1991;
- b) the variables (disposable income, etc.) for the tax system 1994;

3) Load the matrix files that allow us to identify the deciles associate with poor (first decile), middle (from second to ninth decile), and rich (tenth decile) of the distribution of disposable income computed according to 1994 tax system.

4) For each observation and each alternative compute a matrix Fh ($n_record \times na$) whose elements are:

$$F(h_i) = w_i h_i - T(w_i h_i), \quad i = 1, 2, \dots, 15 \quad (8)$$

where $w_i h_i$ is the hourly wage multiplied by the hours associated to each of the 15 alternatives, $T(\cdot)$ is the tax function for 1994. The choice alternatives are not working ($i = 0$), working in the public sector at different hours ($i = 2, 3, 4, 5, 6, 7, 8$) and in the private sector ($i = 9, 10, 11, 12, 13, 14, 15$)

5) Compute a matrix C^* , ($n_record \times na$), whose elements are

$$C^*(h_i, y) = F(h_i) + y \quad (9)$$

where y is a matrix ($n_record \times na$) determined by an iterative procedure so that for $y = \tilde{y}$, at each observation the following equality holds

$$v_i^0 = v_i(F(h_i) + \tilde{y}_i), \quad i = 1, 2, \dots, 15 \quad (10)$$

The value of \tilde{y} is determined using an iterative procedure, described in Appendix B.

4.2. Computation of the Integral

The integral is computed numerically dividing the integration interval in small steps (a length of NOK 100 was found sufficient) and then summing up the partial contribution of the integrand function related to each step. The final result for each observation is obtained by summing the single integrals, evaluated for each alternative, over the total number of alternatives.

4.3. E[CV]

Finally, E[CV] is computed subtracting from I^* (Equation (5)) the integral evaluated in the previous step.

5. Program Structure

The program, whose flow-chart is shown in Appendix C, consists of a main part which resorts to several procedures to accomplish different tasks. They are briefly described below.

Main program:

The main program includes the computation of:

- 1) consumption function using the procedure **V** and **V_SCALAR**
- 2) disposable income computed with:
 - a) woman wage before tax using the procedure **W_WAGE**
 - b) woman net wage using the procedure **NETWAGE_W_94**
 - c) men net wage when woman is not working using the procedure **NETWAGE_M_94**
 - d) men net wage when woman is working using the procedure **NETWAGE_MW_94**
- 3) \tilde{y} : monetary value to be added or subtracted to the

¹Of course these data sets must be previously prepared with the specific program that considers the different tax systems to be estimated.

woman net wage of 1994 (Fh 94) for which the utility of 1994 equals that of 1991. It is evaluated resorting to the procedure INCR and the command lines listed in the main program for the refinement of the interval were the solution lies.

4) the integral using the procedure INTEGRAND.

5) E[CV] and E[CV] statistics for the total sample and for deciles of disposable income distribution (poor (first decile), middle (from second to ninth decile), and rich (tenth decile)).

Procedures (listed in alphabetic order):

HALF: solution refinement through half interval search

INCR: iterate until an interval is found with function values of different sign

INTEGRAND: computation of the function to be integrated and the integral (Equation (5)), for given y , observation i , and alternative j

INTERP: search a solution via linear interpolation

NETWAGE_M_94: compute man net wage when woman is not working (1994 tax system).

NETWAGE_MW_94: compute man net wage, when woman works (1994 tax system).

NETWAGE_W_94: calculation of woman net wage (*i.e.* women working) using tax function 1994

V: computation of consumption function

The procedure returns the matrix v (of dimension $n_record \times na$) of the consumption function for a given matrix of disposable income $disp$ (of dimension $n_record \times na$), according to the following equation:

$$\begin{aligned}
 v_{i,j} = \exp & \left[1.77 \left(\frac{[10^{-4}(C_{i,j} - 60000)]^{0.64} - 1}{0.64} \right) \right. \\
 & + (115.02 - 63.61X_{i1} + 9.20X_{i1}^2 \\
 & + 1.27X_{i2} + 0.97X_{i3}) \left(\frac{\left(1 - \frac{h_j}{3640}\right)^{-0.53} - 1}{-0.53} \right) \\
 & \left. - 0.12 \left(\frac{10^{-4}(C_{i,j} - 60000)^{0.64} - 1}{0.64} \right) \right. \\
 & \left. \times \left(\frac{\left(1 - \frac{h_j}{3640}\right)^{-0.53} - 1}{-0.53} \right) \right] \quad (11)
 \end{aligned}$$

where C_{ij} = disposable income of record i and alternative j , passed as a parameter to the procedure.

The other parameters are passed as global variables and have the following meanings:

X_{i1} is the logarithm of age of the woman, X_{i2} is the number of children aged 0 - 6, and X_{i3} is the number of children aged 7 - 17.

V_SCALAR: computation of the value of consumption function v for a single value of disposable income, a given sample i and alternative j . This is the scalar version of procedure **V**. The procedure returns the value (scalar) of the consumption function v for a given disposable income $disp$, sample i and alternative j according to the equation reported for procedure **V** above.

W_WAGE: Woman wage income before tax.

6. Results

The expected CV obtained from the above procedure is reported in the following **Table 1**.

From **Table 1**, we observe that the mean household in the sample gained NOK 27078 from the 1992 tax reform. The richest household gained almost 10 times more than the poorest or 4 times more in relative income terms.

The distribution of expected gain across households is given in **Figure 1**, and we observe that most of the households will benefit from the 1992 tax reform. Thus, such a reform would have attained support from a clear majority of households with married and cohabiting women at an election.

We have also calculated the expected value of compensating variation of a flat tax reform. In the calculations, the tax-revenue-neutral flat tax reform of 29% is used as a reference. Negative values mean that the numerical values have to be subtracted from household incomes under the flat tax regime in order to make the households indifferent in welfare terms between the 1994 regime and the flat tax regime. The expected CV is reported in the following **Table 2**. This Table then says that, on average, the households will gain NOK 51528 if there is a shift from the 1994 tax regime to a flat tax regime.

The richest households gain around 8 times more than the poorest. Thus, in a distributional sense, the richest household benefited more from having the 1991 regime replaced with the 1994 tax regime than they would have in the case of a shift from the 1994 tax regime to a flat tax regime. In **Figure 2**, we show the population density of the individual mean CV. We observe that a vast majority will benefit from the replacement of the 1994 tax regime with a flat tax regime.

More details of the results are reported in [1].

Table 1. Expected value of compensating variation (in NOK 1994) for the 1992 tax reform. The 1991 tax system is used as a reference against the 1994 tax system.

	E[CV]	E[CV] in percent of observed disposable income*
All	27,078	11.46
Deciles in the distribution of household disposable income*:		
1 (poor)	6,761	4.32
2–9 (middle)	24,896	11.11
10 (rich)	64,150	16.66

*Decile(s) refers to the deciles in the distribution of disposable income, 1994.

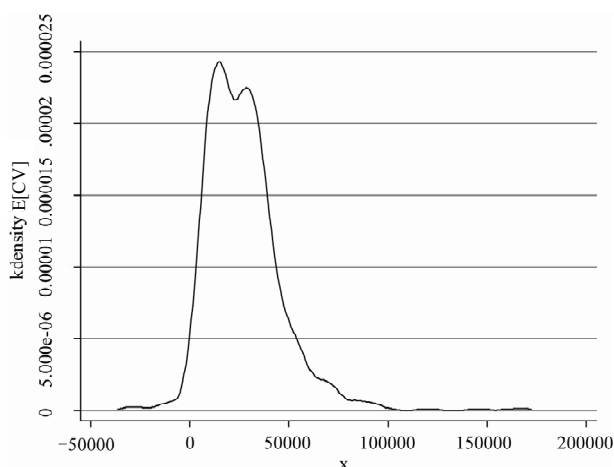


Figure 1. Population density of expected Compensating Variation. Distribution of E[CV], comparing the 1991 tax regime against the 1994 tax regime.

Table 2. Expected value of compensating variation (in NOK 1994) for a flat tax reform. A flat tax regime is used as a reference against the 1994 tax system.

	E[CV]
All	-51,437
Deciles in the distribution of household disposable income, flat tax:	
1 (poor)	-17,155
2–9 (middle)	-53,093
10 (rich)	-146,966

7. Conclusions

In this note we have described a GAUSS software instrument, to compute the value of expected compen-

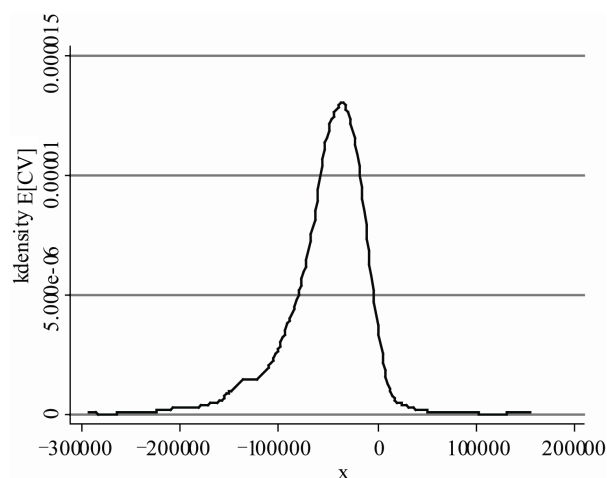


Figure 2. Population density of expected Compensating Variation. Distribution of E[CV], with the flat tax system of 29% used as a reference against the 1994 tax regime.

sating variation within the discrete choice setting suggested in [1].

The program refers to the following tax reforms: a) tax systems in force in 1991 and in 1994 using 1991 as reference year, b) 1994 tax system and a flat tax system taking 1994 as reference year. The program is flexible and can be easily modified to take into account different tax systems and different reference years.

The results show the different impact of the tax reform:

1) from the 1992 tax reform we observe that most of the households will benefit from the 1992 tax reform; the richest household gained almost 10 times more than the poorest or 4 times more in relative income terms. Thus, such a reform would have attained support from a clear majority of households with married and cohabiting women at an election.

2) from the 1994 tax system and a flat tax system taking 1994 as reference year we observe that the richest households gain around 8 times more than the poorest. We observe that a vast majority would benefit from the replacement of the 1994 tax regime with a flat tax regime.

8. Acknowledgements

M. Locatelli gratefully acknowledges financial support by Frisch Centre, Oslo, Norway.

9. References

- [1] J. K. Dagsvik, M. Locatelli and S. Strøm, "Tax Reform, Sector-Specific Labor Supply and Welfare Effect," *Scandinavian Journal of Economics*, Vol. 111, No. 2, 2009, pp. 265-287. [doi:10.1111/j.1467-9442.2009.01565.x](https://doi.org/10.1111/j.1467-9442.2009.01565.x)

- [2] M. Locatelli and S. Strøm, "Computation of the Compensating Variation within a Random Utility Model Using GAUSS Software," Child Working Paper No. 02/2006.
- [3] J. K. Dagsvik and A. Karlstrøm, "Compensating Variation and Hicksian Choice Probabilities in Random Utility Models That are Non-Linear in Income," *Review of Economic Studies*, Vol. 72, No. 1, 2005, pp. 57-76. [doi:10.1111/0034-6527.00324](https://doi.org/10.1111/0034-6527.00324)
- [4] Aptech System, Inc., "GAUSS Software (Version 10)," Black Diamond, Ridgefield, 2009.
- [5] J. K. Dagsvik and S. Strøm, "Sectoral Labor Supply, Choice Restrictions and Functional Form," *Journal of Applied Econometrics*, Vol. 21, No. 6, 2005, pp. 803-826. [doi:10.1002/jae.866](https://doi.org/10.1002/jae.866)
- [6] S. S. Kuo, "Computer Applications of Numerical Methods," Addison-Wesley, Boston, 1972.

Appendix A

Table A1. Estimation results for the parameters of the labor supply probabilities.

Uniformly distributed offered hours with part-time and fulltime peaks			
<i>Variables</i>	<i>Parameters</i>	<i>Estimate</i>	<i>t-values</i>
Preferences:			
<i>Consumption:</i>			
Exponent	α_1	0.64	7.6
Scale 10^{-4}	α_2	1.77	4.2
Subsistence level C_0 in NOK per year		60 000	
<i>Leisure:</i>			
Exponent	α_3	-0.53	-2.1
Constant	α_4	115.02	3.2
Log age	α_5	-63.61	-3.2
(log age) ²	α_6	9.20	3.3
# children 0 - 6	α_7	1.27	4.0
# children 7 - 17	α_8	0.97	4.1
<i>Consumption and Leisure, interaction</i>			
Subsistence level of leisure in hours per year		5120	
The parameters b_1 and b_2; $\log b_j = f_{j1} + f_{j2}S$			
Constant, public sector (sector 1)	f_{11}	-4.20	-4.7
Constant, private sector (sector 2)	f_{21}	1.14	1.0
Education, public sector (sector 1)	f_{12}	0.22	2.9
Education, private sector (sector 2)	f_{22}	-0.34	-3.3
Opportunity density of Offered hours, $g_{k2}(h)$, $k = 1,2$			
Full-time peak, public sector (sector 1)*	$\log(g_1(h_{Full})/g_1(h_0))$	1.58	11.8
Full-time peak, private sector (sector 2)	$\log(g_2(h_{Full})/g_2(h_0))$	1.06	7.4
Part-time peak, public Sector	$\log(g_1(h_{Part})/g_1(h_0))$	0.68	4.4
Part-time peak, private Sector	$\log(g_2(h_{Part})/g_2(h_0))$	0.80	5.2
# observations			810
Log likelihood			-1760.9

Appendix B

To determine the value \tilde{y} we must solve the following equation (for each record and each alternative) $v_0 = v(Fh94 + \tilde{y})$.

That means that we must find the zero of the function $f(x)$ defined as:

$$f(x) = v_0 - v(x), \text{ with } x = Fh94 + y,$$

where y is a generic amount of income to be added to the 1994 woman net wage.

Calling x^* the value for which $f(x) = 0$, we have

$$\tilde{y} = x^* - Fh94$$

To determine the value of x^* the following steps are done:

- 1) iterate until an interval is found where the solution lies, using procedure INCR;
- 2) refine the interval iterating until an approximate solution is found. For the very first iterations (NITER \leq 5), the new value is searched using linear interpolation (see **Figure 1**), implemented by the procedure INTERP. Then (NITER $>$ 5) the solution is refined using a half-interval search method, implemented by the procedure HALF. For more details on these methods see, for example, Ch. 6 of [6].

The exit test is performed only after the solution has been refined using HALF (i.e. only if NITER $>$ 5) and is

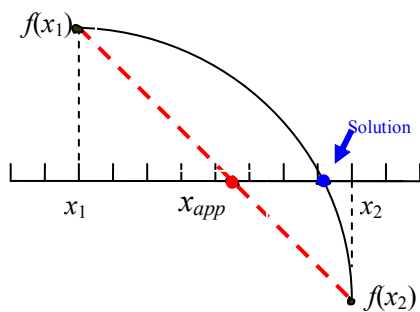


Figure 1. The linear interpolation method: $x_{app} = x_1 +$

$$\frac{f(x_1)(x_2 - x_1)}{f(x_1) - f(x_2)}$$

based both on an absolute and relative tolerance.

Considering the last interval x_1, x_2 where the solution is sought, the function value $f(x_m)$ at the mean value $x_m = 0.5(x_1 + x_2)$ is computed. Furthermore the relative error on x , $rel_err = |x_2 - x_1|/x_1$, is evaluated. The solution is accepted if $|f(x_m)| \leq abs_tol$ or $rel_err \leq rel_tol$.

A satisfactory trade-off between speed and accuracy has been found assuming $abs_tol = 10$ and $rel_tol = 1e-8$. Then the procedure exits assuming $x^* = x_m$.

Appendix C: the GAUSS Program Flow

