TECHNICAL REPORT 2011-1

Calculating Disposable Income as a Function of Primary Income**

Huiling Zhang*#
Osman Icoz*
Richard Arnott*

July 2011

Abstract: The computer program calculates equilibrium rents by property type and wages by labor type (the rental rate on capital is exogenous) for each model zone and time period. There are different groups of individuals. An individual in a particular group is assigned a particular endowment of property by type, labor by type, and capital. From this, in each time period each individual's primary income (PI) is calculated. But because of taxes and transfers, an individual's disposable income (DI), the amount she has to spend (the program allows for no personal saving) in that period, differs from her primary income. This technical report documents how the function that relates disposable income to primary income (DI (PI)) was calculated.

** This report was prepared in conjunction with the University of California MRPI Project LA-Plan.
* Department of Economics, University of California, Riverside, CA 92506#
# Please address questions and comments to the primary author, Huiling Zhang, at hzhan017@ucr.edu.

Calculating Disposable Income as a Function of Primary Income

The computer program calculates equilibrium rents by property type and wages by labor type (the rental rate on capital is exogenous) for each model zone and time period. There are different groups of individuals. An individual in a particular group is assigned a particular endowment of property by type, labor by type, and capital. From this, in each time period each individual's primary income (PI), the income from primary factors of production, is calculated. But because of taxes and transfers, an individual's disposable income (DI), the amount she has to spend (the program allows for no personal saving) in that period, differs from her primary income. This technical report documents how the function that relates disposable income to primary income (DI (PI)) was estimated.

The estimation is non-trivial since there appear to be no readily available data that relate disposable income to primary income. This report has four sections. The first describes the data that were used in the estimation, the second the conceptual basis for the estimation, the third the estimation, and the fourth a concluding section.

With more time and effort, superior estimation procedures could no doubt be devised. This technical report aims only to come up with a broadly reasonable procedure.

1. Data

Two datasets were used to get DI(PI) mapping. They are indexed as dataset A and B hereafter.

A. Internal Revenue Service (IRS) Data Book, 2000. Washington, D.C. [1]

Two tables from the Data Book were employed:

A1. "Tax Year 2000, California Selected Income and Tax Items for Individual Income Tax Returns: Forms 1040, 1040A & 1040EZ, By Size of Adjusted Gross Income."

California data on the following variables were extracted, and are displayed in Table 2 of Appendix I. Note that the unit is a tax return, not distinguished according to whether the return is a single return or a joint return.

(i) Adjusted Gross Income (AGI)

(ii) Taxable Transfers (specifically, Social Security Income (SS), and Unemployment Insurance Compensation (UIC))

(iii) Income tax payable (T) (both federal income tax, and state and local income tax)

A2: "Tax Year 2000, Selected Income and Tax Items for Individual Income Tax Returns: Forms 1040, 1040A & 1040EZ, By Accumulated Size of Adjusted Gross Income."

This table is attached as Table 3 in Appendix I. It was used to estimate the distribution of adjusted gross income for US nationwide tax return filers.

B. "US Treasury Distributional Analysis Methodology", Cronin, J. A. 1999. [2]

Only two tables of that paper are used. Footnote 1 in Table 10 gives data related to the distribution of Family Economic Income (FEI) in US national scale, while Tables 8 and 10 permit estimating a mapping from Total Transfer Income (TTI) to Family Economic Income (FEI) in US national scale.

Those two tables are displayed in Appendix II, and associated calculations for estimations are explained step by step in Appendix IV.

Notice that in Cronin's paper, all the data for year 2000 were estimated data. Therefore, the FEI distribution and TTI(FEI) mapping in this technical report were estimated from estimated data.

2. The Conceptual Basis for the Estimation

The basic procedure is to estimate both primary income and disposable income as functions of adjusted gross income, from which the function relating disposable income to primary income can be obtained.

2.1 Notation

Table 1: Notations

Notation	Economic variable
AGI	adjusted gross income
ATI	after-tax income
DI	disposable income
FEI	family economic income
NTT	non-taxable transfer income
PI	primary income
SS	social security income
T	income tax payable
TTI	total transfer income
UIC	unemployment insurance compensation

2.2 Equations

$$PI(AGI) = AGI - SS(AGI) - UIC(AGI) \quad (1)$$

Primary income equals adjusted gross income less taxable transfer income, defined as social security income plus less unemployment insurance income. SS and UIC are treated as transfer income, rather than as a regular annuity or as actuarially-fair insurance payouts respectively, because both SS and UIC programs are so strongly redistributive in character.

$$ATI(AGI) = AGI - T(AGI) \quad (2)$$

The tax payable includes both federal and state income taxes.

$$TTI(AGI) = TTI(FEI(AGI)) \quad (3)$$

B provides data on the basis of which FEI can be estimated as a function of AGI, and TTI can be estimated as a function of FEI, from which TTI can be estimated as a function of AGI. Details of the procedure are given in the next section.

$$NTT(AGI) = TTI(AGI) - SS(AGI) - UIC(AGI) \quad (4)$$

This equation is simply a definition. Non-taxable transfer income is simply total transfer income minus taxable transfer income.

$$DI(AGI) = ATI(AGI) + NTT(AGI) \quad (5)$$

Disposable income is defined as primary income plus transfers minus taxes. Adjusted gross income equals primary income plus taxable transfers. ATI therefore gives primary income plus taxable transfers, net of tax. Disposable income equals primary income plus taxable transfers, net of tax, plus non-taxable transfers. Thus, (5) is essentially a definition.

$$DI(PI) = DI(AGI(PI)) \quad (6)$$

Knowing both DI and PI as functions of AGI permits the calculation of DI as a function of PI.

3. Estimation Procedure

3.1 Outline

Step 1: Estimate two cumulative distribution functions: G(AGI) and F(FEI)

Step 2: Assume $G(AGI) = F(FEI)$, then derive the mapping FEI(AGI)

Step 3: Estimate three functions PI(AGI)*, T(AGI)*, & TTI(FEI)

Step 4: Use estimated TTI(FEI) to derive TTI(AGI)

Step 5: Derive DI(AGI)

$$\begin{aligned} DI(AGI) &= ATI(AGI) + NTT(AGI) \\ &= AGI - T(AGI) + TTI(AGI) - SS(AGI) - UIC(AGI) \\ &= PI(AGI) + TTI(AGI) - T(AGI) \end{aligned}$$

Step 6: Derive DI(PI)

DI(PI) = DI(AGI(PI)) where AGI(PI) is just the inverse of the PI(AGI).

To sum up, five estimations were conducted. They are two distributions, G(AGI) and F(FEI), and three mappings, PI(AGI)*[1], T(AGI)*, and TTI(FEI). All the other functional relationships can be derived once those five estimations are done.

Among the five estimations, PI(AGI)* and T(AGI)* were based on California data. US national data were used in the other three estimations as a result of lack of California data.

[1] Estimations with asterisk (*) means California data based.

3.2 Step 1&2: Estimate two cumulative distribution functions: G(AGI) and F(FEI). Estimate mapping FEI(AGI)

FEI[2] is income for a family[3], but it also includes singles that are not attached to any family. AGI is income for a tax-return filer, and it contains both joint filers and non-joint filers. Most joint filers are couples within a household, and most non-joint filers are single. Because of the similarity of the units in the two databases, the following assumption was made:

Assumption: *A certain economic unit's rankings in AGI distribution and in FEI distribution (both are for the entire US) are the same. i.e. G(AGI) = F(FEI).*

Then FEI (AGI) can be derived once cdf's of both FEI and AGI are estimated.

A rule of thumb is that income distributions are *lognormal and Pareto in the right tail.* The AGI data contains negative income. To deal with this, the AGI cdf curve

[2] *Family Economic Income*: a family's annual income is the amount it can spend during the year and still have the same net assets at the end of the period as it did at the beginning.

[3] *Family*: Treasury's family includes the taxpayer, and his or her spouse and dependents (if any) living in the same household. A family may include two or more income tax filing units. All families are included, whether or not any member of the family files an income tax return. Dependents, along with their income, are attached to their larger family unit. Single (non-dependent) individuals are considered one-person families. The economic incomes of all members of a family are added together to compute the family's FEI.

More information on definitions and measures of Family and Family Economic Income can be found in Cronin, 1999[2].

was shifted to the right to ensure that the lognormal cdf is applicable. The same procedure was applied to deal with negative FEI incomes. To estimate five parameter values of cdf of FEI given seven data points, a visual iterative procedure had to be used. But for cdf of AGI, piecewise-nonlinear estimation was used on thirty-five data points. Figure 1 and Figure 2 below show the fitted cdf curves for FEI and AGI respectively. Figure 3 and Figure 4 show the fitted cdf curves for (FEI + c1) and (AGI + c2) in logarithmic scale, where c1 and c2 are the translation parameters.

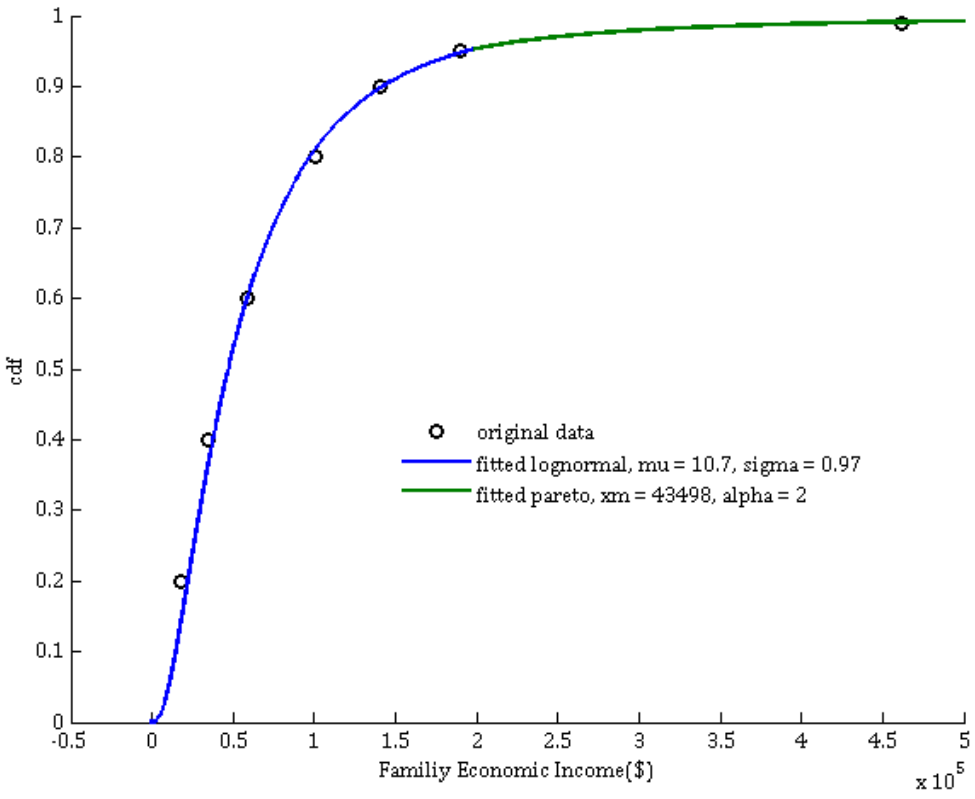

Figure 1: Fitted cdf of US national family economic income (FEI). The distribution of (FEI + c1) is lognormal if (FEI + c1) < $20000, and pareto otherwise, where translation parameter c1 = 2000. Visual iterative procedure was employed in estimation due to lack of data.

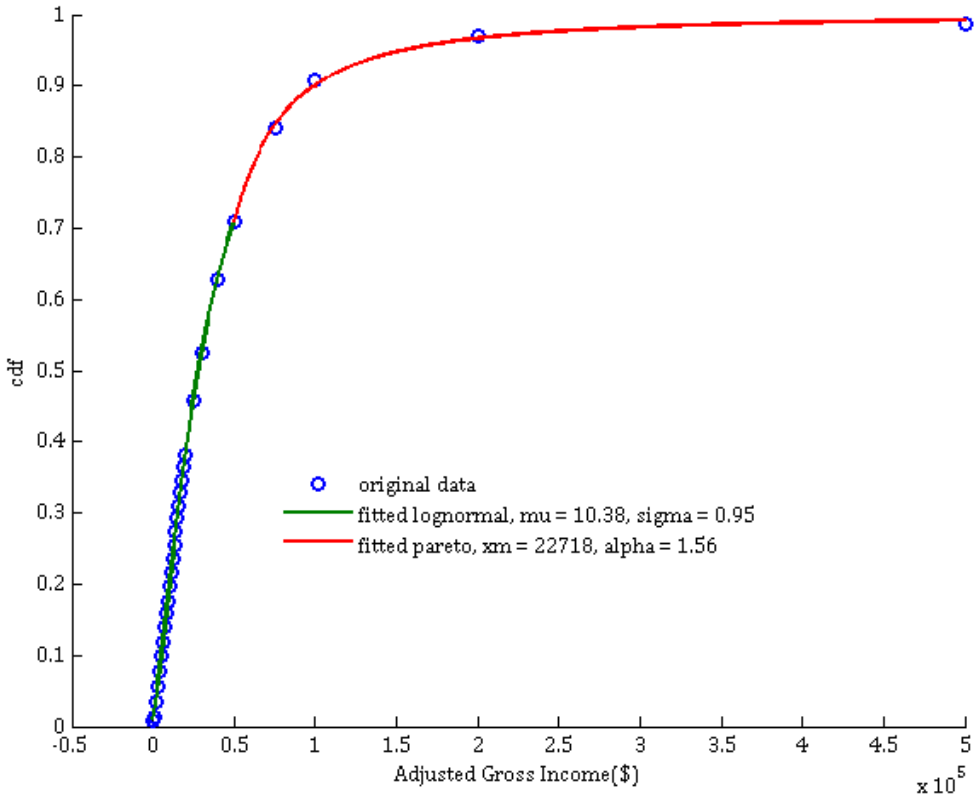

Figure 2: Fitted cdf of US national adjusted gross income (AGI). The distribution of (AGI + c2) is lognormal if AGI < $50000, and pareto otherwise, where translation parameter c2 = 4195. Piecewise-nonlinear estimation method was employed given relatively large data size.

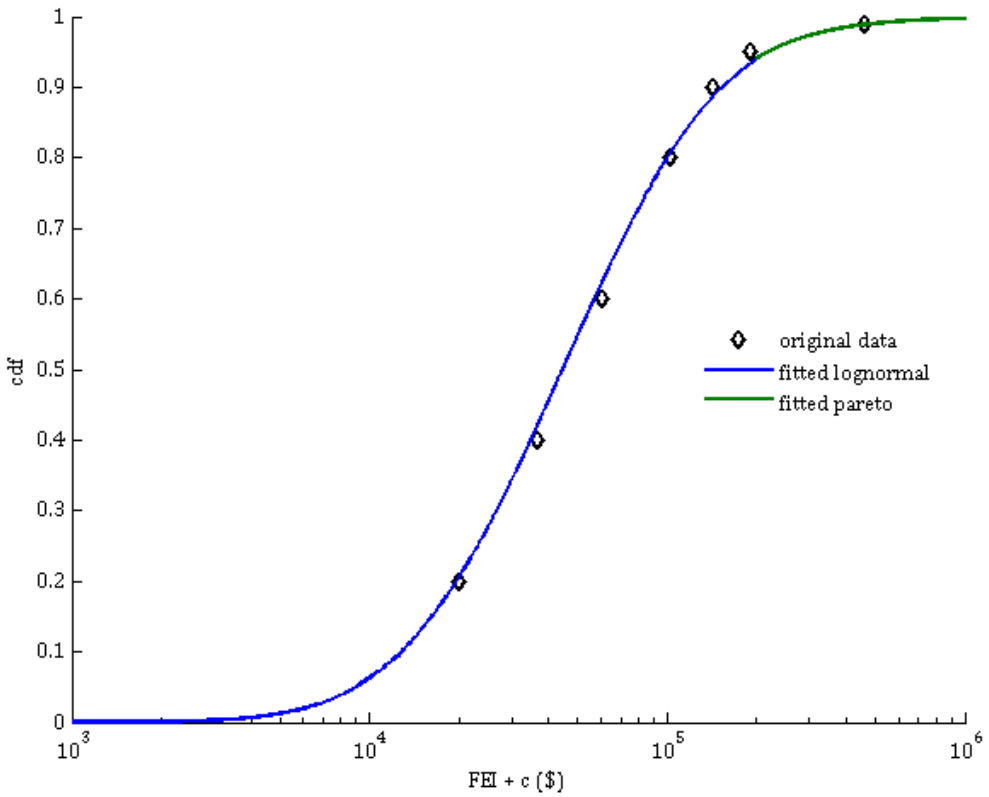

Figure 3: Fitted cdf of family economic income (FEI) after translation. The translation parameter c1 = 2000. The horizontal axis is US national FEI data after transformation (FEI + c1), and is plotted on a logarithmic scale. Visual iterative procedure was employed in estimation due to lack of data.

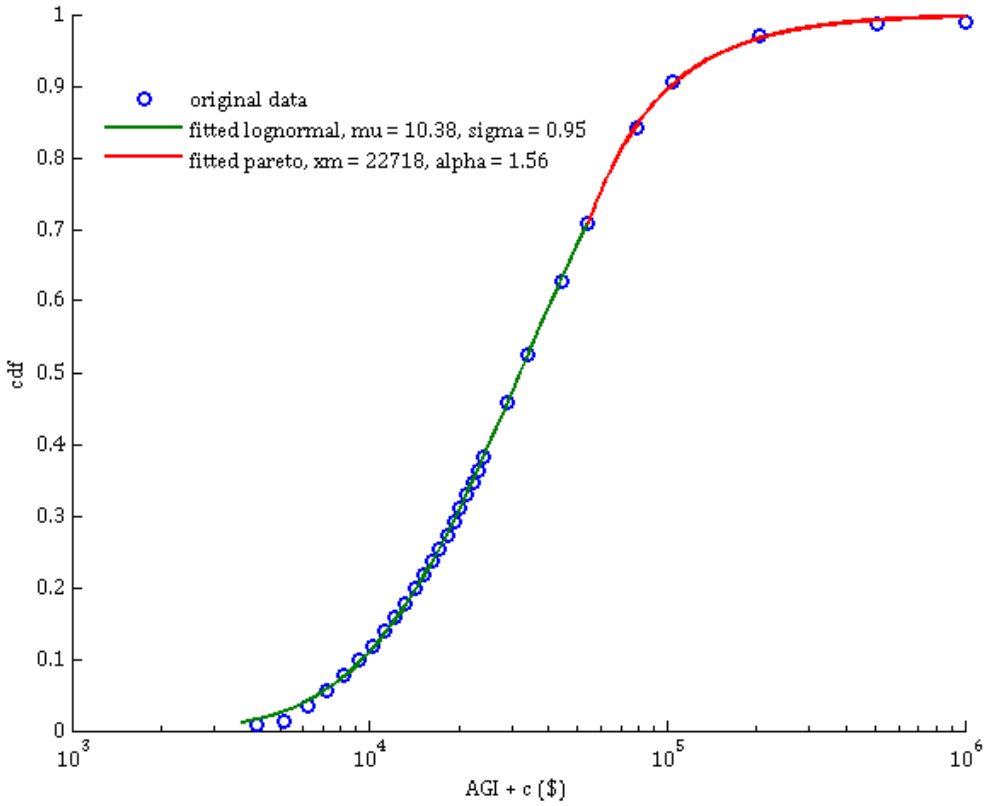

Figure 4: Fitted cdf of US national adjusted gross income (AGI) after translation. The translation parameter c2 = 4195. The horizontal axis is US national AGI data after transformation (AGI + c2), and is plotted on a logarithmic scale. Piecewise-nonlinear estimation method was employed given relatively large data size.

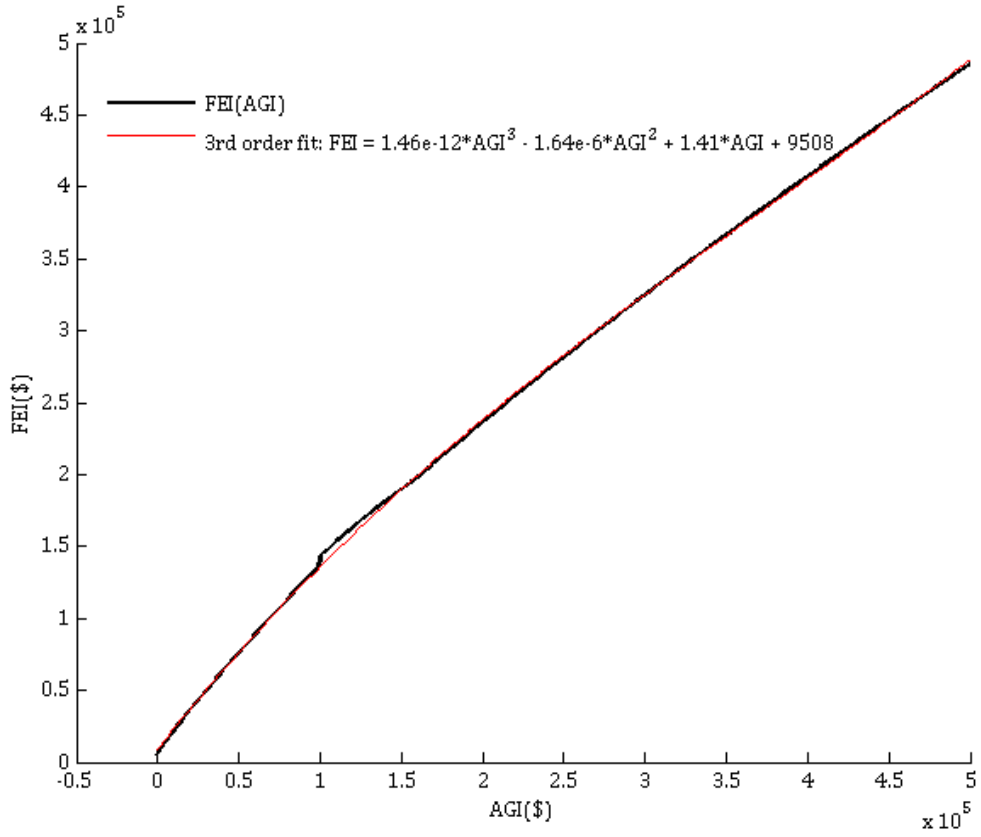

Figure 5: Family economic income (FEI) as a function of adjusted gross income (AGI) for US national. This function is derived from equation G(AGI) = F(FEI), where G(AGI) and F(FEI) are the estimated cdf's of AGI and FEI respectively from US national data. The derived FEI(AGI) function should be considerably complicated, but the plot fits well into a third order polynomial:

$$FEI = 1.46e_{-12} \times AGI^3 - 1.64e_{-6} \times AGI^2 + 1.41 \times AGI + 9508$$

3.3 Step 3: estimate three functions: PI(AGI)*, T(AGI)*, & TTI(FEI)

3.3.1 PI(AGI)*

The Data Book provided twelve intervals for AGI. The data points below in Figure 8 correspond to the mean values for each of these intervals. Since the plot of PI against AGI is very close to a straight line, PI(AGI)* was estimated using linear least squares regression. Details of the procedure are given in Table 4.1 of Appendix III.

Explicit function forms of mapping UIC(AGI)* and mapping SS(AGI)* are not necessary in estimating PI(AGI)*, therefore Figure 6 and Figure 7 are displayed here to convey more information about data used in this work.

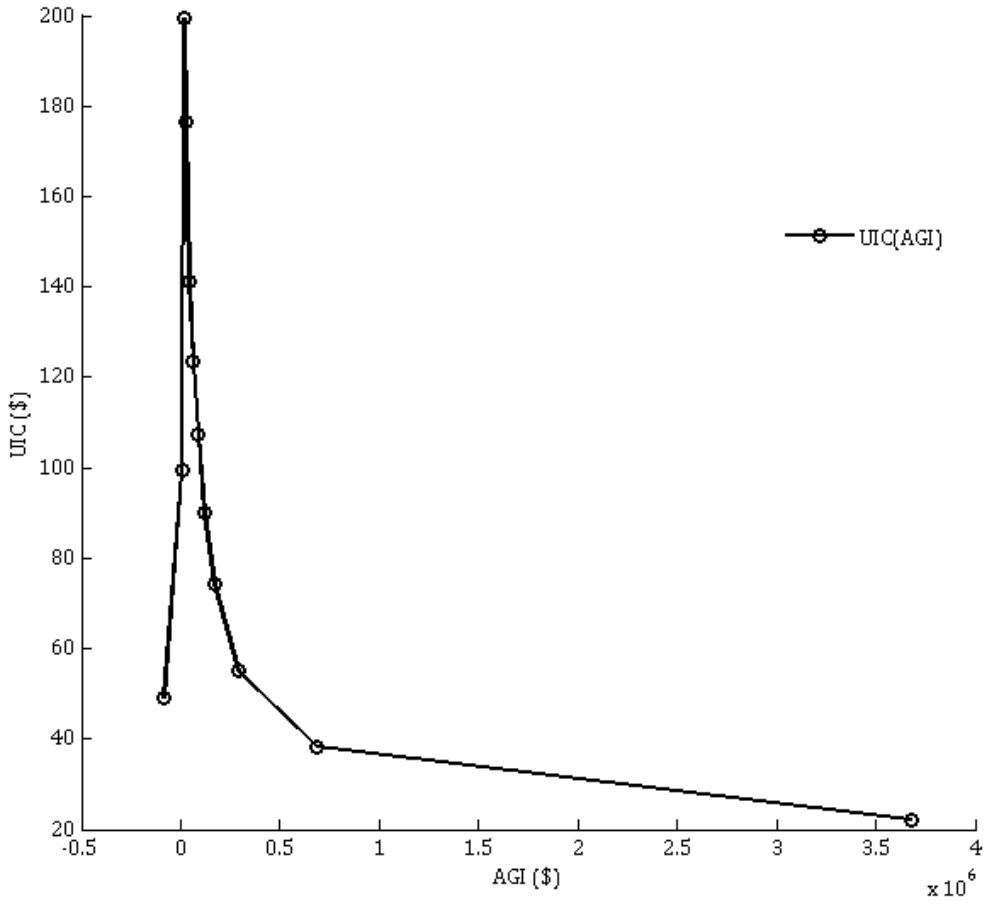

Figure 6: Unemployment insurance compensation (UIC) versus adjusted gross income (AGI) for California data. Since UIC(AGI) function is not necessary for deriving PI(AGI) function, the mapping is not estimated.

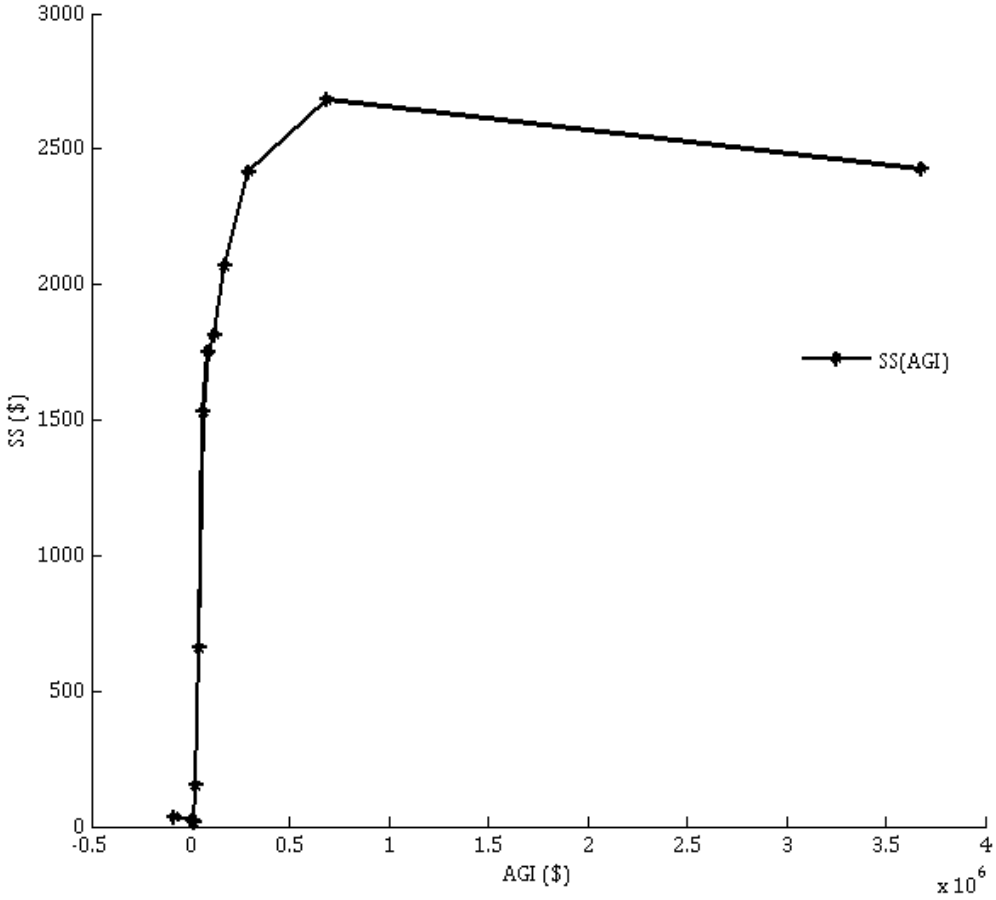

Figure 7: Social security income (SS) versus adjusted gross income (AGI) for California data. Since SS(AGI) function is not necessary for deriving PI(AGI) function, the mapping is not estimated.

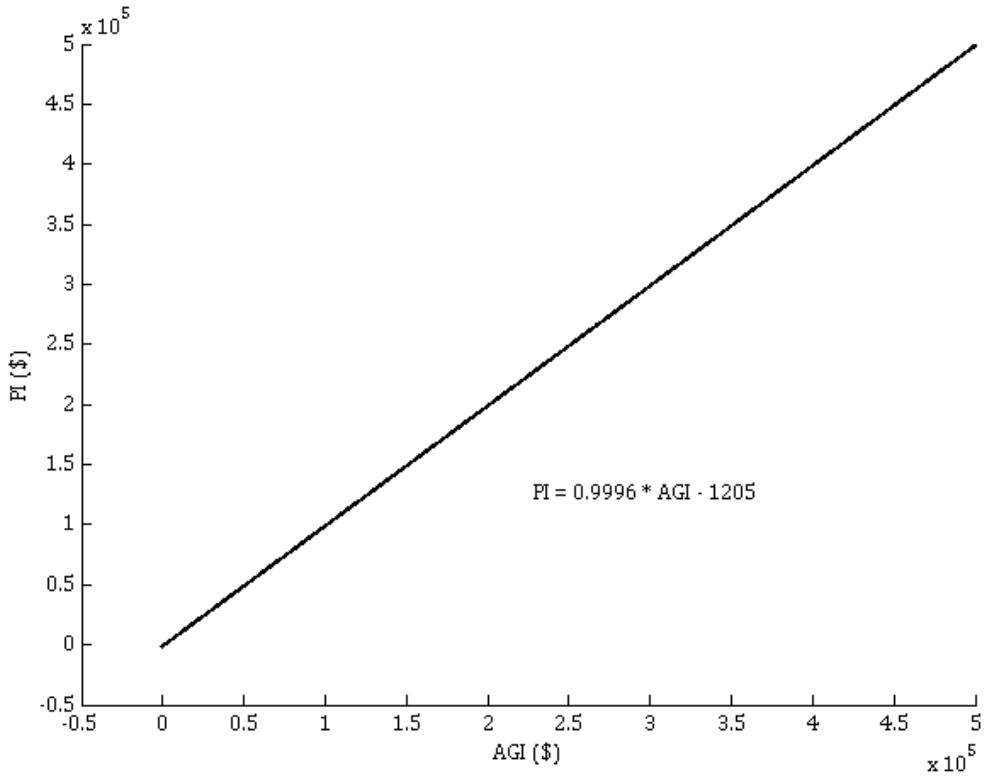

Figure 8: Primary income (PI) was estimated as a linear function of adjusted gross income (AGI) for California data. PI = 0.9996 * AGI - 1205

3.3.2 T(AGI)*

Twelve points are available for estimating, for CA residents, Income Tax payable (federal + state) as a function of Adjusted Gross Income. Again, interval means are used. Since the point for the highest AGI is far away from the others, so that considerable information between the second highest and the highest AGI interval is missing, several curves fit the data equally well. Economic common sense is used as the selection criterion. Details of the procedure are given in Table 4.3 of Appendix III.

The curve was estimated to be made up of two pieces: low AGI piece is 3rd order polynomial and high AGI piece is linear, as shown in Figure 9.

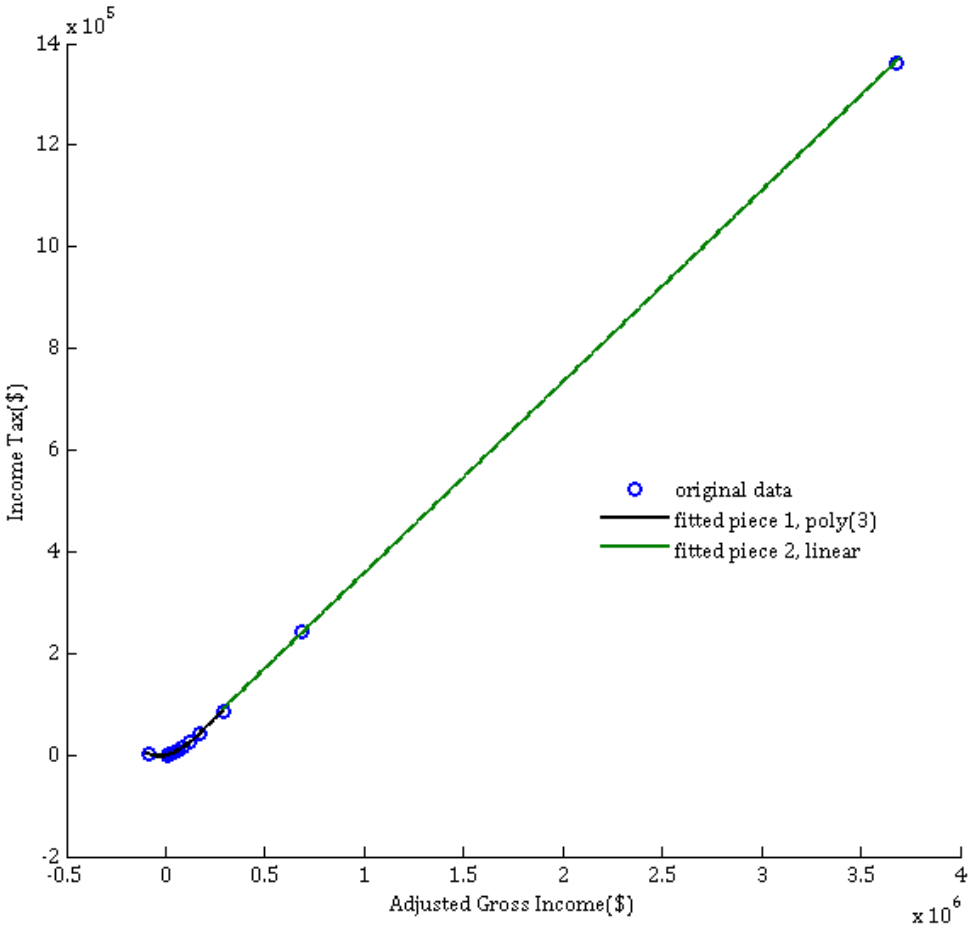

Figure 9: Income tax payable (T) estimated as a function of adjusted gross income (AGI) for California data. The function has two pieces: one is third order polynomial, and the other is linear.

$$T = (1.7179e-12) \times AGI^3 + (1.2545e-6) \times AGI^2 + 0.0852 \times AGI - 929, \quad if\ AGI < 8.9414e4;$$
$$= 0.3759 \times AGI - 1.8123e4, \quad otherwise$$

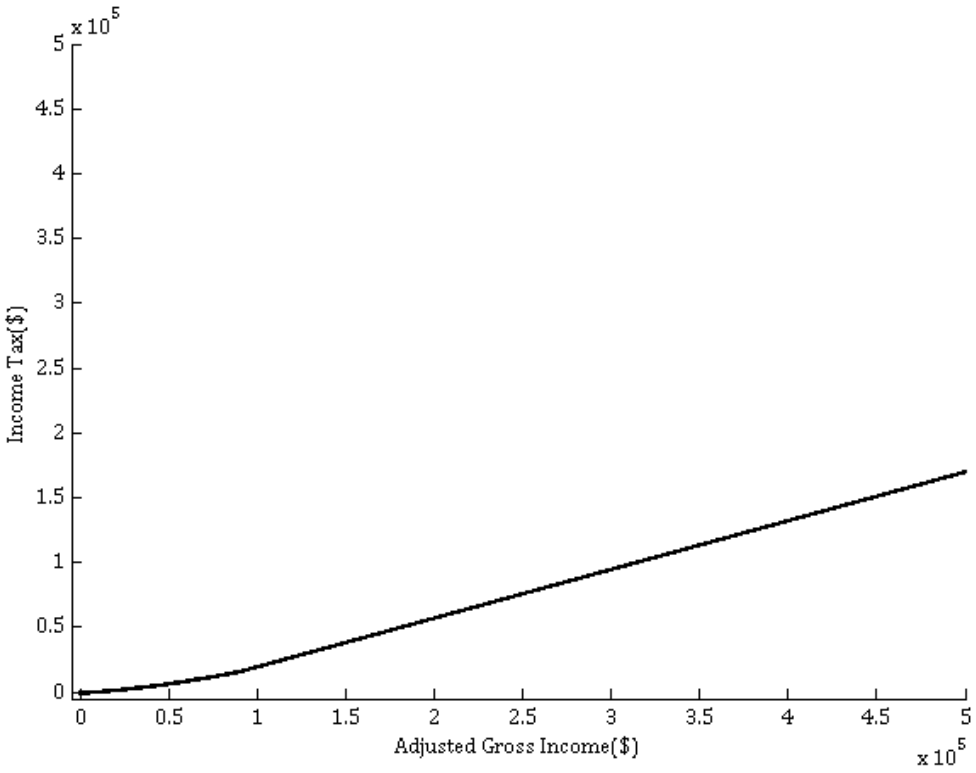

Figure 10: Income tax payable (T) estimated as a function of adjusted gross income (AGI) for California data. Data with AGI up to 4,000,000$ were used in estimation, but this figure only displays the less than 500,000$ part, which is of major interest to the project.

3.3.3 TTI(FEI)[4]

This is the least reliable of all the estimations. The data used are average FEI for a family by FEI quintile versus the average TTI received by a family in that quintile. They are group average data, so they are very likely to be biased. Group average data were also used to estimate PI(AGI)*, but since the function was approximately linear, the group average data are almost unbiased. However, TTI(FEI) function here is non-linear non-monotonic, and its shape does not conform to any of the standard

[4] Transfer income includes nontaxable Social Security benefits, Supplemental Security Income (SSI), Temporary Assistance for Needy Families (TANF), general assistance, other support, food stamps, Low Income Housing Energy Assistance (LIHEA), veterans' compensation, and workers' compensation.
More information on definitions and measures of Total Transfer can be found in Cronin, 1999[2].

fitting functions. In the end, the curve was fitted with four linear segments as shown in Figure 11.

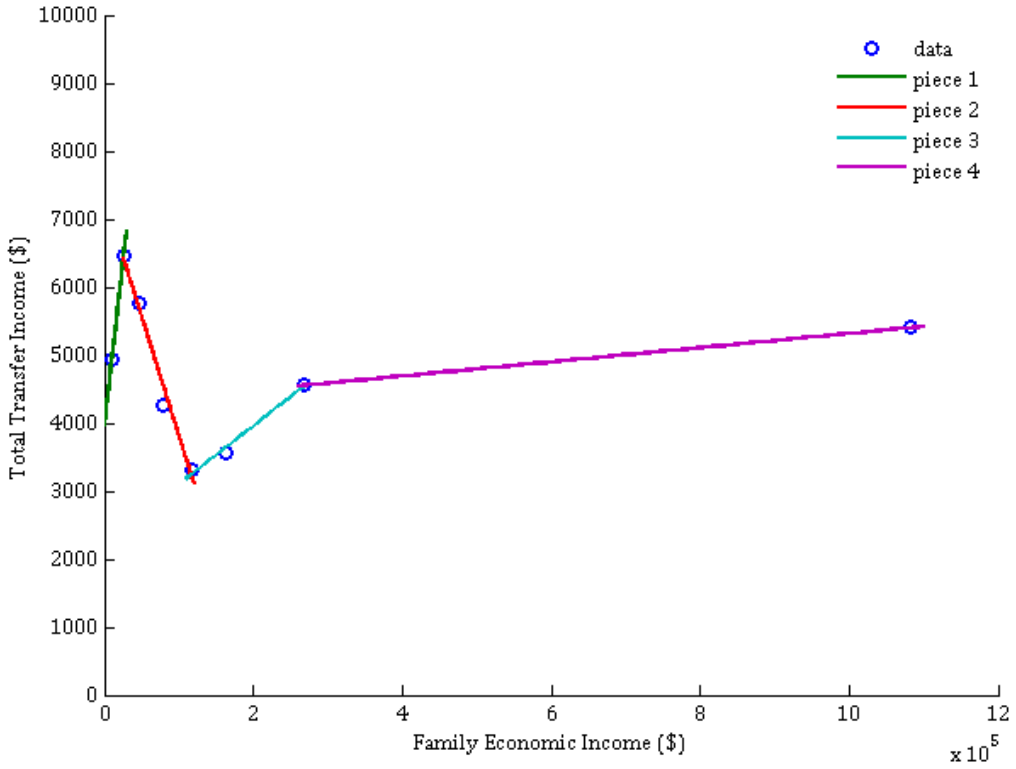

Figure 11: total transfer income (TTI) as a function of family economic income (FEI). The function is piecewise linear. $TTI = 0.0962 \times FEI + 3973, \quad if \quad FEI < 2.5e4$
$= -0.0352 \times FEI + 7328, \quad if \quad 2.5e4 < FEI < 1.2e5$
$= 0.0086 \times FEI + 2248, \quad if \quad 1.2e5 < FEI < 2.7e5$
$= 0.0010 \times FEI + 4283, \quad if \quad 2.7e5 < FEI$

3.4 After Step 3:

Step 4, 5, and 6 are straightforward, entailing simply algebraic manipulation of the functions estimated above. Figure 12 below displays the relationship between DI and PI. Remarkably, the relationship is almost linear. Fitting a linear function use ordinary least squares gives

$$DI = 0.6264 \times PI - 529 \qquad (7)$$

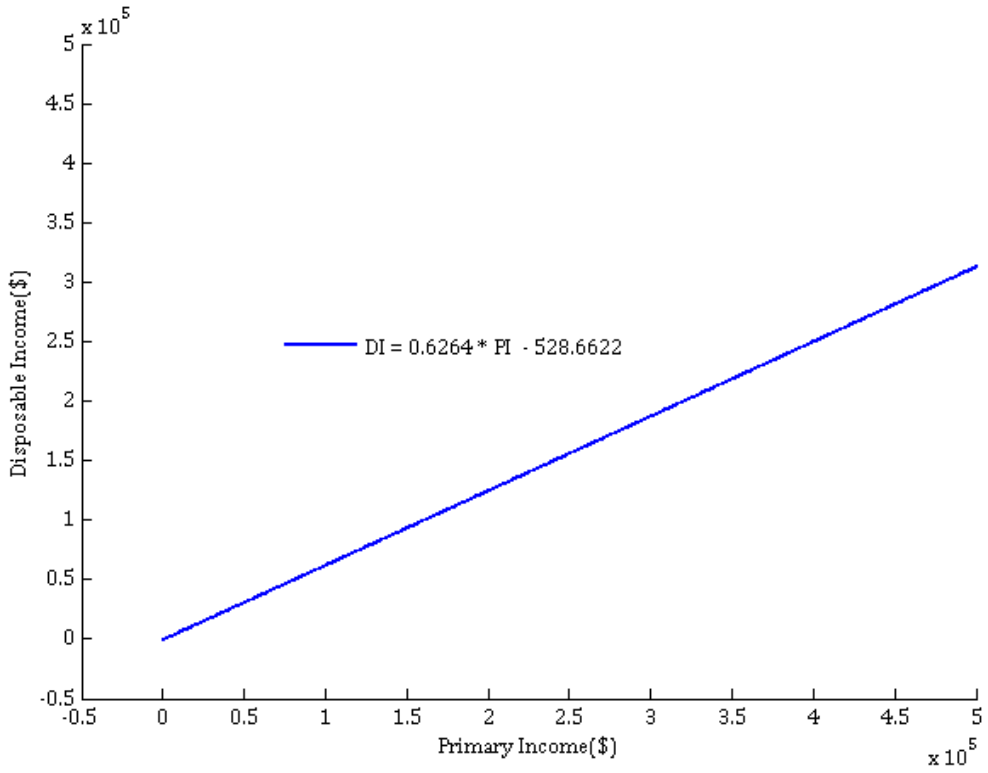

Figure 12: disposable income (DI) estimated as a linear function of primary income (PI).

4. Concluding Comments

The computer program calculates the *primary income* of each economic unit. An economic unit's primary income is calculated as its endowments of the various types of labor multiplied by the corresponding (endogenous) equilibrium wage rates, plus its endowment of capital times the exogenous rental rate on capital, plus its endowment of various types of property times the corresponding (endogenous) asset values times the exogenous competitive rate of return. But the amount of income an economic unit has to spend, its *disposable income*, is different from this due to taxes and transfers. This report has documented the procedure followed in estimating the relationship between primary income and disposable income for tax

filers in California in the year 2000. Remarkably, the relationship was estimated to be a straight line, with a slope of 0.6264.

The estimation was done by applying several sensible yet heroic assumptions applied to fragmentary data. The rationales for the assumptions made are given in the body of the report. Here some additional qualifications and caveats are added.

1. The procedure treated social security payments as taxable transfers on the grounds that the social security program is strongly redistributive. It did not however treat social security taxes. This should be done by applying a social security tax function to an economic unit's wage income generated by the computer.

2. The retail sales tax should be treated by distinguishing between producer prices and consumer prices. Producer prices equal unit costs. Consumer prices equal producer prices times the sales tax rate. Consumer prices enter the economic units' budget constraints.

3. No attention was paid to corporate income taxes. In public economic theory, the corporate income tax is treated as a mix of a tax on pure profits and a tax on income generated in the corporate sector. In the theory underlying the model, there are no pure profits in equilibrium. Thus, for conceptual consistency, the corporate income tax should be regarded as driving a wedge between the producer and "consumer" (factor supplier) rental rate on capital. Capital is typically assumed to be freely mobile. Consistency with this assumption entails treating the consumer rental rate on capital as exogenous, with the producer rental rate equaling the

exogenous rental rate on capital times one plus the effective rate on capital. This will need to be estimated and incorporated into the computer program.

4. The report took the economic unit to be an income tax filer, and did not distinguish between joint and single returns. The computer program however takes individuals as the economic unit. It remains to be determined how these individuals will be defined. Originally, the aim was to determine the relationship between disposable income and primary income for individuals. This turned out to be impractical since the IRS Data Book did not provide separate data for single and joint filers. Fortunately, the issue appears moot. Since the estimation indicates that the *net* of transfers marginal income tax rate is constant over the entire income range, the same tax rate can be applied to individuals, however defined.

The output of the computer program includes per individual primary income by year and by demographic group. An individual's consumption however is based on disposable income, which is primary income less taxes plus transfers. The goal of this technical report was, on the basis of empirical data on the current income tax and transfer system in California, to estimate the relationship between individual disposable income and individual primary income. Due to a lack of data at the individual level, the relationship between *household* disposable income in 2000 (DI) and *household* primary income in 2000 (PI) was estimated. Furthermore, because data were unavailable by demographic, an aggregated relationship was estimated, the same for all demographic groups. The estimated relationship was DI = 0.6264*PI

– 528.6622. To avoid spurious accuracy, it is recommended that the following formula be used in the program for the year 2000:

$$DI = 0.625*PI - 500.$$

Note that this formula does *not* take into account social security payments (OASDI) or the corporate income tax. Thus, the computer program should add a line adjusting disposable income to account for social security payments. It should also adjust capital income to account for corporate income taxes paid. Also, an adjustment needs to be made to account for the program's using individuals rather than households as the economic unit. The sales tax should be taken into account by distinguishing between the consumer and producer prices of goods and services. Finally, consideration needs to be given concerning whether the function should be modified for years after 2000.

REFERENCES

1. Internal Revenue Service (IRS) Data Book 2000. Washington, D.C. (http://www.irs.gov/taxstats/indtaxstats/article/0,,id=98123,00.html)

2. Cronin, J. A. (1999), U.S. Treasury Distributional Analysis Methodology (Office of Tax Analysis, Washington, DC).

Appendix I: Raw data from dataset A1.

Two tables were extracted from IRS Data Book. They are listed here as Table 2 and Table 3.

Table 2 has data on California income tax return filers. It is useful in estimating PI(AGI)* and T(AGI)*.

Table 3 has data on US nationwide income tax filers. The first and third column permits estimating distribution of AGI.

Table 2: Data Extracted from IRS (xxxx) and Calculations

California, 2000 (Amounts are in thousands of dollars)	Total	Size of Adjusted Gross Income											
		Breakeven and Loss	**$0.01 Under $10,000**	**$10,000 Under $20,000**	**$20,000 Under $30,000**	**$30,000 Under $50,000**	**$50,000 Under $75,000**	**$75,000 Under $100,000**	**$100,000 Under $150,000**	**$150,000 Under $200,000**	**$200,000 Under $500,000**	**$500,000 Under $1,000,000**	**$1,000,000 and Over**
Returns Count(#)	14,866,950	157,868	2,675,592	2,745,921	2,029,500	2,695,844	1,929,657	1,040,628	853,508	298,048	331,473	63,540	45,371
cdf of AGI		0.0106	0.1906	0.3753	0.5118	0.6931	0.8229	0.8929	0.9503	0.9704	0.9927	0.9969	1.0000
AGI Amount	864,644,512	-13,206,174	14,230,183	40,767,433	50,209,204	105,348,523	118,259,374	89,700,491	102,454,317	51,086,868	95,734,902	43,317,987	166,741,404
Average AGI		-83.65	5.32	14.85	24.74	39.08	61.29	86.20	120.04	171.40	288.82	681.74	3675.07
federal income tax	146,453,548	16,956	220,817	1,603,424	3,133,889	9,327,002	12,967,125	11,628,966	15,976,354	9,362,526	21,983,873	11,930,829	48,301,788
FIT average		0.11	0.08	0.58	1.54	3.46	6.72	11.17	18.72	31.41	66.32	187.77	1064.60
State and Local Income Tax:	40,953,147	245,261	48,802	277,800	280,661	1,455,971	3,129,764	3,447,164	5,153,710	3,103,488	6,825,191	3,454,231	13,531,103
SLIT average		1.55	0.02	0.10	0.14	0.54	1.62	3.31	6.04	10.41	20.59	54.36	298.23
Average F & S income tax (T)		1.66	0.10	0.69	1.68	4.00	8.34	14.49	24.76	41.83	86.91	242.13	1,362.83
Unemployment Compensation:													
Number of Returns	894,111	2,677	141,018	234,526	146,821	167,338	105,766	48,434	31,518	8,520	6,560	629	304
Amount	2,031,107	7,734	266,255	547,781	357,882	381,004	238,026	111,570	77,041	22,124	18,256	2,429	1,006
Average UIC		0.05	0.10	0.20	0.18	0.14	0.12	0.11	0.09	0.07	0.06	0.04	0.02
Taxable Social Security Income:													
Number of Returns	1,157,534	759	4,160	19,781	162,398	316,318	279,809	146,406	115,731	43,103	52,329	10,410	6,330
Amount	10,206,036	5,436	58,935	39,285	307,320	1,776,119	2,950,052	1,824,228	1,546,887	616,419	800,897	170,346	110,111
Average SS		0.03	0.02	0.01	0.15	0.66	1.53	1.75	1.81	2.07	2.42	2.68	2.43
PI average		-83.74	5.20	14.63	24.41	38.28	59.63	84.34	118.14	169.26	286.35	679.02	3,672.62

Note: **Black** rows of data are extracted from original source, while **blue** rows are calculated from the black rows.

Individual Income Tax Returns, Tax Year 2000

Table 3 --Selected Income and Tax Items, by Size and Accumulated Size of Adjusted Gross Income--Continued

(All figures are estimates based on samples--money amounts are in thousands of dollars)

Size and accumulated size of adjusted gross income	All returns					Taxable returns			
	Number of returns	Percent of total	Adjusted gross income less deficit			Number of returns	Percent of total	Adjusted gross income less deficit	
			Amount	Percent of total	Average (dollars)			Amount	Percent of total
	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)
Accumulated from Smallest Size of Adjusted Gross Income									
No adjusted gross income	1,146,357	0.9	-58,599,965	[1]	-51,118	5,714	[2]	-5,646,432	[1]
$1 under $1,000	1,819,159	1.4	1,070,919	[2]	589	190,266	0.2	161,154	[2]
$1 under $2,000	4,423,774	3.4	4,991,602	0.1	1,128	721,350	0.7	948,930	[2]
$1 under $3,000	7,203,752	5.6	11,918,609	0.2	1,655	1,136,415	1.2	1,992,705	[2]
$1 under $4,000	10,084,661	7.8	21,980,768	0.3	2,180	1,493,487	1.5	3,232,222	0.1
$1 under $5,000	12,802,742	9.9	34,203,382	0.5	2,672	2,299,944	2.4	6,962,943	0.1
$1 under $6,000	15,389,476	11.9	48,443,225	0.8	3,148	3,260,951	3.4	12,252,604	0.2
$1 under $7,000	17,985,543	13.9	65,350,771	1.0	3,634	4,032,211	4.2	17,266,088	0.3
$1 under $8,000	20,568,744	15.9	84,705,415	1.3	4,118	5,116,888	5.3	25,420,880	0.4
$1 under $9,000	22,957,578	17.7	104,999,828	1.6	4,574	6,369,331	6.6	36,095,685	0.6
$1 under $10,000	25,604,343	19.8	130,179,042	2.0	5,084	7,831,447	8.1	49,980,764	0.8
$1 under $11,000	28,046,437	21.7	155,793,402	2.4	5,555	9,156,493	9.5	63,883,927	1.0
$1 under $12,000	30,538,788	23.6	184,451,536	2.9	6,040	10,528,376	10.9	79,659,636	1.3
$1 under $13,000	32,905,745	25.4	214,066,004	3.3	6,505	11,856,278	12.2	96,270,115	1.6
$1 under $14,000	35,307,245	27.3	246,507,155	3.8	6,982	13,244,144	13.7	115,026,911	1.9
$1 under $15,000	37,715,408	29.2	281,422,506	4.4	7,462	14,759,276	15.2	137,007,773	2.3
$1 under $16,000	40,152,140	31.0	319,191,125	5.0	7,950	16,296,097	16.8	160,838,017	2.6
$1 under $17,000	42,500,226	32.9	357,915,803	5.6	8,422	17,791,857	18.4	185,501,521	3.0
$1 under $18,000	44,794,654	34.6	398,052,630	6.2	8,886	19,281,946	19.9	211,572,458	3.5
$1 under $19,000	47,133,195	36.4	441,314,446	6.9	9,363	20,813,638	21.5	239,914,293	3.9
$1 under $20,000	49,376,942	38.2	485,024,222	7.6	9,823	22,307,472	23.0	269,018,662	4.4

$1 under $25,000	59,370,057	45.9	709,413,488	11.0	11,949	29,873,324	30.9	439,508,222	7.2
$1 under $30,000	67,738,815	52.4	938,789,229	14.6	13,859	37,170,515	38.4	639,920,813	10.5
$1 under $40,000	81,286,542	62.8	1,409,682,178	21.9	17,342	50,172,997	51.8	1,092,347,262	17.9
$1 under $50,000	91,698,632	70.9	1,875,285,627	29.2	20,451	60,423,249	62.4	1,550,813,502	25.5
$1 under $75,000	108,774,791	84.1	2,919,940,682	45.5	26,844	77,384,113	79.9	2,588,768,058	42.5
$1 under $100,000	117,372,119	90.7	3,657,444,293	56.9	31,161	85,964,771	88.8	3,324,842,995	54.6
$1 under $200,000	125,455,566	97.0	4,723,786,040	73.5	37,653	94,042,640	97.1	4,390,471,817	72.1
$1 under $500,000	127,591,329	98.6	5,337,541,678	83.1	41,833	96,176,686	99.3	5,003,726,485	82.2
$1 under $1,000,000	127,987,460	98.9	5,606,562,565	87.3	43,806	96,572,441	99.7	5,272,491,266	86.6
$1 under $1,500,000	128,086,969	99.0	5,727,166,792	89.2	44,713	96,671,856	99.8	5,392,978,910	88.6
$1 under $2,000,000	128,131,551	99.0	5,803,877,628	90.3	45,296	96,716,398	99.9	5,469,621,571	89.8
$1 under $5,000,000	128,198,319	99.1	6,003,271,106	93.5	46,828	96,783,091	100.0	5,668,792,957	93.1
$1 under $10,000,000	128,215,928	99.1	6,123,848,481	95.3	47,762	96,800,683	100.0	5,789,250,278	95.1
$1 or more	128,227,143	99.1	6,423,976,613	100.0	50,098	96,811,889	100.0	6,088,909,264	100.0
All returns	**129,373,500**	**100.0**	**6,365,376,648**	**99.1**	**49,202**	**96,817,603**	**100.0**	**6,083,262,832**	**99.9**

Appendix II: Raw data on FEI and TTI.

This appendix reproduces the Tables that were drawn from Cronin, J. A. (1999). Footnote 1 in Table 10 gives data related to the distribution of Family Economic Income (FEI), while Tables 8 and 10 permit estimating a mapping from Total Transfer Income (TTI) to Family Economic Income (FEI).

Table 10: Percentage Distribution of FEI Factor Incomes in 2000

	Labor Income		Capital Income		Transfer Income		FEI
Family Economic Income Quintile	As a percent of Total FEI	Percent Distribution	As a percent of Total FEI	Percent Distribution	As a percent of Total FEI	Percent Distribution	Percent Distribution
Lowest[1]	44.9	1.7	6.1	0.7	49.0	19.2	2.7
Second	62.3	6.3	13.0	4.2	24.8	25.8	7.2
Third	71.4	12.7	16.1	9.2	12.5	23.0	12.6
Fourth	78.9	23.6	15.6	15.1	5.5	17.0	21.3
Highest	69.7	55.6	28.5	73.2	1.8	14.8	56.7
Total[1]	71.1	100.0	22.1	100.0	6.9	100.0	100.0
Top 10%	65.0	37.0	33.6	61.5	1.4	8.3	40.5
Top 5%	60.1	24.9	38.8	51.8	1.1	4.8	29.4
Top 1%	52.3	10.9	47.1	31.6	0.5	1.2	14.8

1. Families with negative incomes are excluded from the lowest quintile but included in the total line. Quintiles begin at FEI of: Second $17,988; Third $34,844; Fourth $59,019; Highest $100,767; Top 10% $140,581; Top 5% $189,835; Top 1% $462,053.

Table 8: Distribution of Family Economic Income (FEI) at 1996 and 2000 levels[1]

Family Economic Income Quintile	1996 levels			2000 levels		
	Families (millions)	Family Economic Income ($B)	Family Economic Income (%)	Families (millions)	Family Economic Income ($B)	Family Economic Income (%)
Lowest[2]	21.4	185	2.9	22.4	226	2.7
Second	21.9	492	7.8	23.0	602	7.2
Third	21.9	844	13.3	23.0	1,062	12.6
Fourth	21.9	1,364	21.5	23.0	1,790	21.3
Highest	21.9	3,473	54.9	23.0	4,771	56.7
Total[2]	109.4	6,330	100.0	115.2	8,419	100.0
Top 10%	10.9	2,465	38.9	11.5	3,407	40.5
Top 5%	5.5	1,787	28.2	5.8	2,480	29.5
Top 1%	1.1	905	14.3	1.2	1,247	14.8

1. The total FEI levels for the two years should not be construed as representing growth in FEI over this period; the definition of FEI is not the same for the two periods. The 1996 level figures do not include the employer share of payroll taxes, nor do they adjust for state income tax refunds and the self-employed health deduction, and the 1996 unreported income amounts are not targeted to the National Income and Product Accounts (see Section 5.2).
2. Families with negative incomes are excluded from the lowest quintile but included in the total line. Quintiles at 1996 levels begin at FEI of: Second $15,604; Third $29,717; Fourth $48,660; Highest $79,056; Top 10% $108,704; Top 5% $145,412; Top 1% $349,438. Quintiles at 2000 levels begin at FEI of: Second $17,988; Third $34,844; Fourth $59,019; Highest $100,767; Top 10% $140,581; Top 5% $189,835; Top 1% $462,053.

Appendix III: Three estimations based on data source A.

Data source A was useful in three estimations:

(i) cdf of Adjusted Gross Income (AGI); (ii) Primary Income(PI) as a function of Adjusted Gross Income(AGI); and (iii) income tax payable (T) as a function of Adjusted Gross Income (AGI).

This appendix documents how data needed in three estimations were calculated.

Table 4.1: estimate cdf of Adjusted Gross Income (AGI)

California, 2000 (Amounts are in thousands of dollars)	Total	Size of Adjusted Gross Income ($1,000)											
		Loss and Breakeven	$0.01 under $10	$10 under $20	$20 under $30	$30 under $50	$50 under $75	$75 under $100	$100 under $150	$150 under $200	$200 under $500	$500 under $1,000	$1,000 and Over
Returns Count(#)	14866950	157868	2675592	2745921	2029500	2695844	1929657	1040628	853508	298048	331473	63540	45371
cdf of AGI		0.01	0.19	0.38	0.51	0.69	0.82	0.89	0.95	0.97	0.99	1.00	1.00
AGI Amount	864644512	-13206174	14230183	40767433	50209204	105348523	118259374	89700491	102454317	51086868	95734902	43317987	166741404
Average AGI		-83.65	5.32	14.85	24.74	39.08	61.29	86.20	120.04	171.40	288.82	681.74	3675.07
federal income tax	146453548	16956	220817	1603424	3133889	9327002	12967125	11628966	15976354	9362526	21983873	11930829	48301788
FIT average		0.11	0.08	0.58	1.54	3.46	6.72	11.17	18.72	31.41	66.32	187.77	1064.60
State and Local Income Tax:	40953147	245261	48802	277800	280661	1455971	3129764	3447164	5153710	3103488	6825191	3454231	13531103
SLIT average		1.55	0.02	0.10	0.14	0.54	1.62	3.31	6.04	10.41	20.59	54.36	298.23
Average F & S income tax (T)		1.66	0.10	0.69	1.68	4.00	8.34	14.49	24.76	41.83	86.91	242.13	1362.83
Unemployment Compensation:													
Number of Returns	894111	2677	141018	234526	146821	167338	105766	48434	31518	8520	6560	629	304
Amount	2031107	7734	266255	547781	357882	381004	238026	111570	77041	22124	18256	2429	1006
Average UIC		0.05	0.10	0.20	0.18	0.14	0.12	0.11	0.09	0.07	0.06	0.04	0.02
Taxable Social Security Income:													
Number of Returns	1157534	759	4160	19781	162398	316318	279809	146406	115731	43103	52329	10410	6330
Amount	10206036	5436	58935	39285	307320	1776119	2950052	1824228	1546887	616419	800897	170346	110111
Average SS		0.03	0.02	0.01	0.15	0.66	1.53	1.75	1.81	2.07	2.42	2.68	2.43

1. In the data set, tax-return filers are classified into twelve categories according to the size of their AGI.
   There's no data on the minimum and maximum of AGI, so eleven cut-off points are available. They are {0, 10, 20, 30, 50, 75, 100, 150, 200, 500, and 1000} in thousands of dollars.
2. cdf for a certain cutoff point = the number of tax returns that have AGI amount less than or equal to this cutoff point (#) / total retuns count(#)

Table 4.2 : estimate Primary Income(PI) as a function of Adjusted Gross Income(AGI) [1]

California, 2000 Amounts are in thousands of dollars	Total	Size of Adjusted Gross Income ($1,000) Loss and Breakeven	$0.01 under $10	$10 under $20	$20 under $30	$30 under $50	$50 under $75	$75 under $100	$100 under $150	$150 under $200	$200 under $500	$500 under $1,000	$1,000 and Over
Returns Count(#)	14866950	157868	2675592	2745921	2029500	2695844	1929657	1040628	853508	298048	331473	63540	45371
cdf of AGI		0.011	0.191	0.375	0.512	0.693	0.823	0.893	0.950	0.970	0.993	0.997	1.000
AGI Amount	864644512	-13206174	14230183	40767433	50209204	105348523	118259374	89700491	102454317	51086868	95734902	43317987	166741404
Average AGI		**-83.65**	**5.32**	**14.85**	**24.74**	**39.08**	**61.29**	**86.20**	**120.04**	**171.40**	**288.82**	**681.74**	**3675.07**
federal income tax	146,453,548	16,956	220,817	1,603,424	3,133,889	9,327,002	12,967,125	11,628,966	15,976,354	9,362,526	21,983,873	11,930,829	48,301,788
FIT average		0.11	0.08	0.58	1.54	3.46	6.72	11.17	18.72	31.41	66.32	187.77	1064.60
State and Local Income Tax:	40,953,147	245,261	48,802	277,800	280,661	1,455,971	3,129,764	3,447,164	5,153,710	3,103,488	6,825,191	3,454,231	13,531,103
SLIT average		1.55	0.02	0.10	0.14	0.54	1.62	3.31	6.04	10.41	20.59	54.36	298.23
Average F & S income tax (T)		1.66	0.10	0.69	1.68	4.00	8.34	14.49	24.76	41.83	86.91	242.13	1,362.83
Unemployment Compensation:													
Number of Returns [2]	894111	2677	141018	234526	146821	167338	105766	48434	31518	8520	6560	629	304
Amount	2031107	7734	266255	547781	357882	381004	238026	111570	77041	22124	18256	2429	1006
(-) Average UIC		**0.05**	**0.10**	**0.20**	**0.18**	**0.14**	**0.12**	**0.11**	**0.09**	**0.07**	**0.06**	**0.04**	**0.02**
Taxable Social Security Income:													
Number of Returns [3]	1157534	759	4160	19781	162398	316318	279809	146406	115731	43103	52329	10410	6330
Amount	10206036	5436	58935	39285	307320	1776119	2950052	1824228	1546887	616419	800897	170346	110111
(-) Average SS		**0.03**	**0.02**	**0.01**	**0.15**	**0.66**	**1.53**	**1.75**	**1.81**	**2.07**	**2.42**	**2.68**	**2.43**
(=) Average PI [4]		**-83.74**	**5.20**	**14.63**	**24.41**	**38.28**	**59.63**	**84.34**	**118.14**	**169.26**	**286.35**	**679.02**	**3,672.62**

1. The underlying equation used was : PI = AGI - SS(AGI) - UIC(AGI)
2. The number of returns filed for UIC was less than the number of total tax returns for a certain AGI group. The reason is that not all tax filers were eligible to receive UIC.
   But here : Average UIC for an AGI group = UIC amount for the group / number of total tax returns of the group (including those who did not receive UIC)
   Because in this report, people were different only by their incomes. So by averaging over all tax returns of a group, the UIC benefit an average return filer received was then derived.
3. similarly for Average SS.
4. Twelve pairs of "Average AGI" and "Average PI" were then used to estimate PI(AGI) function

Table 4.3: estimate income tax payable (T) as a function of Adjusted Gross Income (AGI) [1]

California, 2000 Amounts are in thousands of dollars	Total	Size of Adjusted Gross Income ($1,000) Loss and Breakeven	$0.01 under $10	$10 under $20	$20 under $30	$30 under $50	$50 under $75	$75 under $100	$100 under $150	$150 under $200	$200 under $500	$500 under $1,000	$1,000 and Over
Returns Count(#)	14866950	157868	2675592	2745921	2029500	2695844	1929657	1040628	853508	298048	331473	63540	45371
cdf of AGI		0.0106	0.1906	0.3753	0.5118	0.6931	0.8229	0.8929	0.9503	0.9704	0.9927	0.9969	1.0000
AGI Amount	864644512	-13206174	14230183	40767433	50209204	105348523	118259374	89700491	102454317	51086868	95734902	43317987	166741404
Average AGI		**-83.65**	**5.32**	**14.85**	**24.74**	**39.08**	**61.29**	**86.20**	**120.04**	**171.40**	**288.82**	**681.74**	**3675.07**
federal income tax	146453548	16956	220817	1603424	3133889	9327002	12967125	11628966	15976354	9362526	21983873	11930829	48301788
FIT average		0.11	0.08	0.58	1.54	3.46	6.72	11.17	18.72	31.41	66.32	187.77	1064.60
State and Local Income Tax:	40953147	245261	48802	277800	280661	1455971	3129764	3447164	5153710	3103488	6825191	3454231	13531103
SLIT average		1.55	0.02	0.10	0.14	0.54	1.62	3.31	6.04	10.41	20.59	54.36	298.23
Average F & S income tax (T)		**1.66**	**0.10**	**0.69**	**1.68**	**4.00**	**8.34**	**14.49**	**24.76**	**41.83**	**86.91**	**242.13**	**1,362.83**
Unemployment Compensation:													
Number of Returns	894111	2677	141018	234526	146821	167338	105766	48434	31518	8520	6560	629	304
Amount	2031107	7734	266255	547781	357882	381004	238026	111570	77041	22124	18256	2429	1006
Average UIC		0.049	0.100	0.199	0.176	0.141	0.123	0.107	0.090	0.074	0.055	0.038	0.022
Taxable Social Security Income:													
Number of Returns	1157534	759	4160	19781	162398	316318	279809	146406	115731	43103	52329	10410	6330
Amount	10206036	5436	58935	39285	307320	1776119	2950052	1824228	1546887	616419	800897	170346	110111
Average SS		0.03	0.07	0.01	0.15	0.66	1.53	1.75	1.81	2.07	2.42	2.68	2.43

1. Income tax = Federal Income tax + State and Local Income tax

Appendix IV: Estimating TTI (FEI) from data source B.

This appendix documents how data pairs of group average (TTI, FEI) were calculated from tables in Appendix II. Three tables are displayed here to show the calculations step by step.

Four columns[5] were extracted from tables in Appendix II. They constitute the main body of the three tables below. Table 5.1 completes the data on interval total FEI, and Table 5.2 derives the interval average FEI and TTI, and Table 5.3 reconstruct intervals to avoid overlaps.

[5] They are highlighted by a red surrounding circle.

Table 5.1 : Calculate interval total FEI for top 10%, top 5%, and top 1% families

	Number of Families (millions)	FEI percent distributions[2] (%)	Family Economic Income[3] ($Billions)	TTI as a Percent of total FEI (%)
Lowest Quintile	22.4	2.7	226	49
Second Q	23	7.2	602	24.8
Third Q.	23	12.6	1062	12.5
Fourth Q.	23	21.3	1790	5.5
Fifth Q.	23	56.7	4771	1.8
top 10%	11.5	40.5	3400.24[4]	1.4
top 5%	5.75	29.4	2468.32	1.1
top 1%	1.15	14.8	1242.56	0.5

1. data in black are original (from table 8* and table 10*), but data in blue are calculated.
2. interval total FEI as a percentage of national total FEI (denoted as FEI_percent)
3. Total FEI for an interval(denoted as FEI_int)
4. Given FEI percent distributions for the three groups, interval total FEI can be calculated once national FEI is obtained.

*FEI_int = FEI_nat * FEI_percent*

Each one of the five quintiles provides a calculation of national total FEI (FEI_nat), but they took slightly different values. So the average of them was used.

Table 5.2 : Calculate Average TTI and Average FEI for each interval

	Number of Families (millions)	FEI percent distributions (%)	Family Economic Income ($Billions)	TTI as a Percent of total FEI (%)	TTI ($Billions)[1]	Average FEI $(thousands)[2]	Average TTI $(thousands)[3]
Lowest Quintile	22.4	2.7	226	49	110.74	10.09	4.94
Second Q	23	7.2	602	24.8	149.30	26.17	6.49
Third Q.	23	12.6	1062	12.5	132.75	46.17	5.77
Fourth Q.	23	21.3	1790	5.5	98.45	77.83	4.28
Fifth Q.	23	56.7	4771	1.8	85.88	207.43	3.73
top 10%	11.5	40.5	3400.24	1.4	47.60	295.67	4.14
top 5%	5.75	29.4	2468.32	1.1	27.15	429.27	4.72
top 1%	1.15	14.8	1242.56	0.5	6.21	1080.48	5.40

1. TTI = Family Economic Income * TTI as a percentage of total FEI
2. Average FEI = Total FEI of an interval/ number of families in that interval
3. Average TTI = TTI / Number of Families

Table 5.3 : Split the fifth quintile into four non-overlapping intervals

	Number of Families (millions)	FEI percent distributions (%)	Family Economic Income ($Billions)	TTI as a Percent of total FEI (%)	TTI ($Billions)	Average FEI $(thousands)	Average TTI $(thousands)
Lowest Quintile	22.4	2.7	226	49	110.74	10.09	4.94
Second Q	23	7.2	602	24.8	149.30	26.17	6.49
Third Q.	23	12.6	1062	12.5	132.75	46.17	5.77
Fourth Q.	23	21.3	1790	5.5	98.45	77.83	4.28
Fifth Q.	23	56.7	4771	1.8	85.88	207.43	3.73
top 10%	11.5	40.5	3400.24	1.4	47.60	295.67	4.14
top 5%	5.75	29.4	2468.32	1.1	27.15	429.27	4.72
top 1%	1.15	14.8	1242.56	0.5	6.21	1080.48	5.40
top 20% - 10%	11.5		1370.76		38.27	119.20	3.33
top 10% - top 5%	5.75		931.92		20.45	162.07	3.56
top 5% - top 1%	4.6		1225.77		20.94	266.47	4.55

1. Since interval average data of FEI and TTI was used in estimation, it is better to use smaller intervals to reduce estimation bias. Therefore, three existing intervals (fifth Q, top 10%, and top 5%) were reconstructed into three new intervals (top 20% - 10%, top 10% - top 5%, and top 5% - top 1%).
2. The non-overlapping four intervals of the fifth quintile, and the first four quintiles together produce eight pairs of interval average FEI and TTI.