CPS ASEC 2005 Tax Model Documentation

Estimates for tax year 2004 were generated from person records in the CPS 2005 ASEC. The methodology used to produce these estimates was updated in tax year 2003 (2004 ASEC). This release continues the use of the new model. Annual federal and state tax law, parameter and threshold changes are included in the modeling. This document explains the most notable changes for tax year 2004.

1. Changes in tax variable imputations

The most significant change in the 2004 tax model estimates involves the output, not the method, of the imputations used to complete the tax returns created for the CPS records. This statistical match uses the Statistics of Income (SOI) public use file produced by the IRS to obtain six variables[1] not collected in the ASEC. The match assigns a wide range of values and approximates the distribution of the variables in the source data, but with a three-year lag. The most recent SOI file pertaining to tax year 2001 was used to impute the tax variables to CPS tax year 2004. Last year, the match between the 2000 SOI file and CPS tax year 2003 was problematic: the donor data had values for many fields that far exceeded reasonable values for the base data, especially for capital gains. In 2003, the most recent aggregates released by the IRS were used to limit the assignment and amount of capital gains. This resulted in very low levels being assigned to the CPS file. This year, the 2001 SOI file was a better donor file for the 2004 CPS file so fewer adjustments were required to impute the fields. As in the previous year, the 2001 file was ratio- adjusted using published IRS aggregates to age the six variables. **Figure 1** shows the distribution of capital gains imputations by AGI percentile. Tax year 2004 more closely replicates the distribution than tax year 2003. This finding is also seen in **Figure 2**, showing aggregate amounts of capital gains assigned by AGI percentile. **Figure 2** includes information on both the 2000 and 2001 SOI microdata, tax years 2003 and 2004 using the new tax model, and for tax year 2002 using the former tax model. The bold line for tax year 2004 remains consistent with the SOI files, while CPS 2002 and 2003 fall short at the upper end of the income distribution.

Two important changes to capital gains assigned in the match are worth noting: First, the maximum value of capital gains assigned was raised. For tax year 2003,[2] the range of possible records from the SOI was limited based on total reported income. This limited the maximum values of all imputed variables, including capital gains. For tax year 2004, SOI records with capital gains in the top one-half percent were restricted from use as donors. Different thresholds were set for high income returns and for dependent returns. This allowed much higher values to remain eligible for matching purposes than the previous truncation method. **Figure 3** shows that the percentage of

---

[1] Capital gains and losses, IRA contributions, self-employed health insurance expenses, SEP contributions, itemized deductions, and childcare expenses.

[2] The Alternative Income and Experiment Poverty Reports using ASEC 2004 (tax year 2003) also contain values for ASEC 2003 (tax year 2002) using the new tax model for year-over comparison purposes. The official tax estimates available by FTP or Ferret for the ASEC 2003 use the old methodology.

returns with capital gains or losses for CPS tax year 2004 approximates that of the donor microdata, the 2001 SOI for most of the AGI classes.

The second major change for tax year 2004 was the treatment of capital gains at the low end of the income distribution. The SOI population is that of filers. The low end of the income distribution of filers is not the same as the low end of the income distribution of the general population, as many of them have no filing requirement. Many low income persons and households do file to obtain refundable credits, but many who are eligible to do so fail to file. The statistical match blocks both the CPS and SOI files on variables to align them for imputation. Income is one of the blocking variables. Low income records in the SOI file are more likely to have capital gains than low income records in the entire CPS population, and the unconstrained match does not adequately limit the assignment of such matches. Therefore, for tax year 2004, the match was performed and then capital gains were zeroed out for randomly selected filers who had income to poverty ratios below 300%[3]. This reduces the likelihood that gains or losses imputed, not reported, could skew experimental poverty calculations. **Figure 3** shows the result of this correction for the lowest AGI classes.

The amount of capital gains and losses imputed differs substantially between tax years 2003 and 2004 due to the procedures described. **Figure 4** compares the percent of returns by the amount of capital gains imputed between tax years 2003 and 2004. In 2004, more returns were assigned less than $1000 in capital gains, but the percentage of returns with gains less than $5000 was approximately equal between the years. A larger difference is apparent in the higher amounts of gains. In tax year 2004, more returns had large amounts of capital gains assigned.

2. Updates to state tax models

Parameters change annually for the major credits and deductions in the federal tax model, but the same base model is used each year. The same is not true for the state tax models. In these, credit programs may begin or end from year to year. The tax model incorporates as many of these changes as possible. **Table 1** offers an overview of what was modeled for each state with an income tax for tax year 2004.

3. New variable names

Several new variables will be released on the public use extract. These include intermediate values used to create final tax estimates. **Table 2** lists the new variables and, where applicable, how they relate to variables from previous years. Two variables

---

[3] The income to poverty ratio (IPR) was constructed by dividing the annual family income by the relevant official poverty threshold. The ratio was distributed to all members in the family and the value for the estimated primary filer was used. Similar IPRs were developed for the SOI data, using total reported income instead of family income. In SOI 2001, 2.90% of single filers, 2.09% of married joint filers, and 1.13% of head of household filers had capital gains. Capital gains assigned to CPS returns above these percentages were randomly deselected.

are now released for both state and federal taxes, one representing the tax estimate before refundable credits, the other providing the tax estimate after all credits. The advance child tax credit and earned income credit are the refundable credits that separate the two tax amounts. Both the child tax credit and the advance child tax credit are public use variables this year; the former is a non-refundable credit as opposed to the latter.

4. Comparison of income, credits and taxes

Table 3 presents weighted aggregate estimates for tax years 2003 and 2004 and compares them to the most recent aggregate values published by the IRS. **Figure 5** illustrates that the large increase in estimated aggregate federal taxes is concentrated at the top of the AGI distribution.

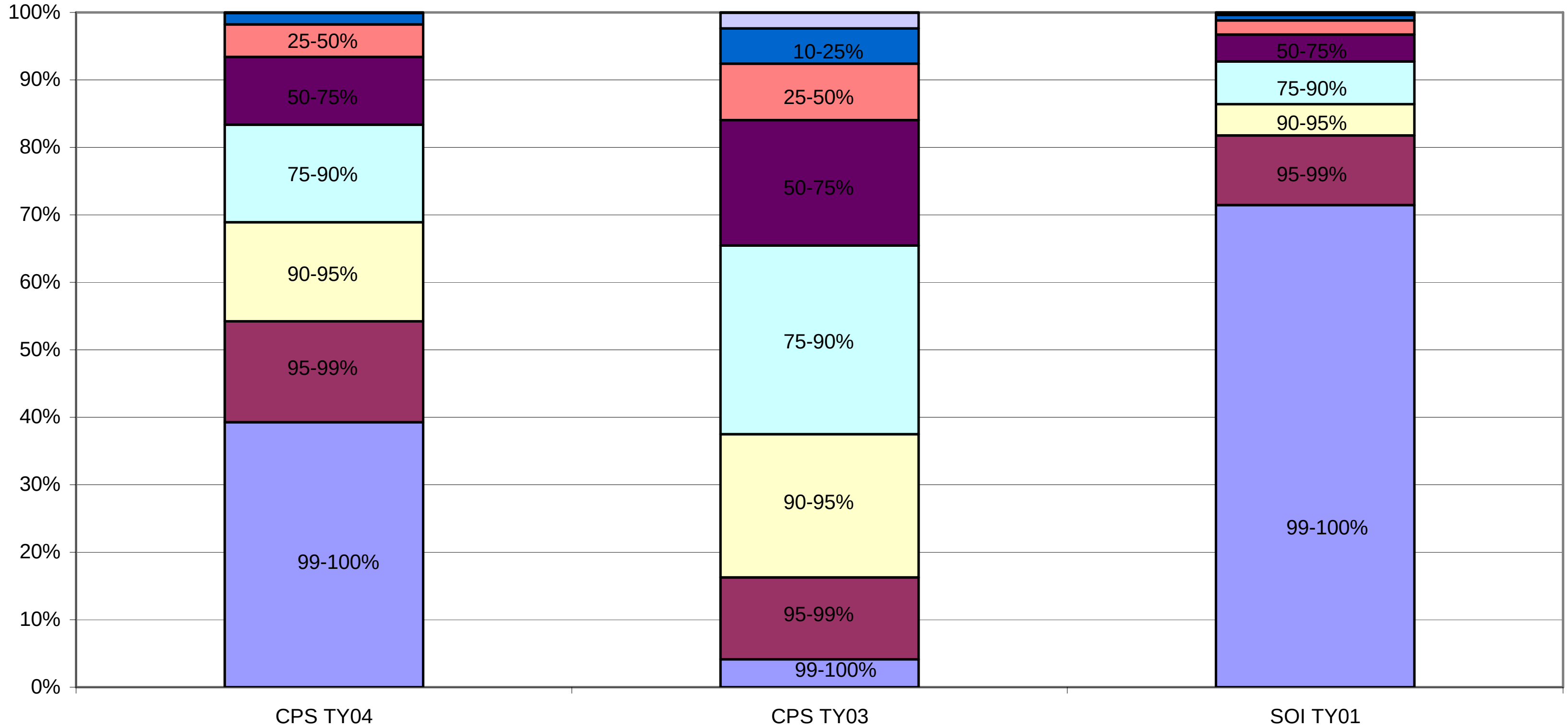
5. Remaining shortcomings of the model

The tax model estimation process is one of continual adjustment. Annual tax law changes, parameters, thresholds and rates are incorporated. The estimates are further evaluated against available statistics to identify areas for improvement. The area receiving most attention this year was the data used in the statistical match. The IRS SOI microdata is a weighted sample of all tax returns with relevant weights. Using the full sample in an unconstrained statistical match can produce unwanted outcomes, with large capital gains or itemized deductions being assigned to a CPS return. An annual review of the SOI is needed given the three-year lag between it, the donor, and the CPS base file. Research is underway to improve or replace this match.

How the estimates rely on the imputed fields is another subject of research. The current imputation assigns total itemized deductions and its components from the donor SOI return. Amounts for fields such as medical and dental expenses and interest expenses are used in other federal schedules or state tax income computations. The effects of the manipulations used to age the lagged data should be evaluated.

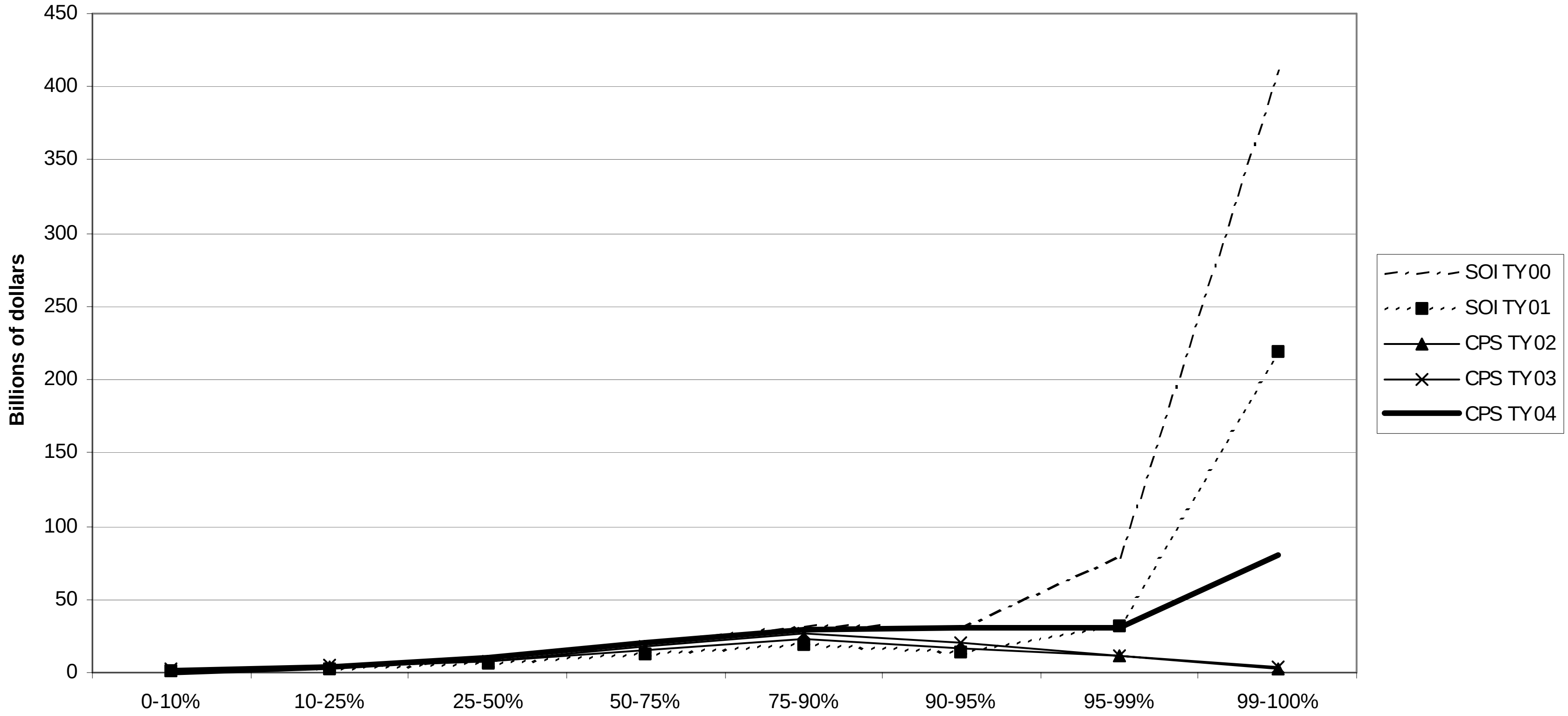
When modeling taxes, the balance between law and likelihood is difficult to address. Currently, the tax model defers to the IRS tax code completely. In reality, tax unit formation may differ substantially if the unit's likely decisions were modeled. This difference would most impact head of household filing.

Figure 1: Distribution of Aggregate Capital Gains by AGI Percentile



Capital gains imputed for the tax year 2004 CPS file (2005 ASEC) are distributed differently than those assigned in tax year 2003 (2004 ASEC). The tax year 2004 distribution better reflects the most recently released IRS microdata from the Statistics of Income public use file, shown on the right. The weighted aggregate capital gains amounts were substantially larger in the CPS ASEC 2005 than in CPS ASEC 2004 – $205 billion versus $97 billion – due to larger gains being assigned to the top of the income distribution.

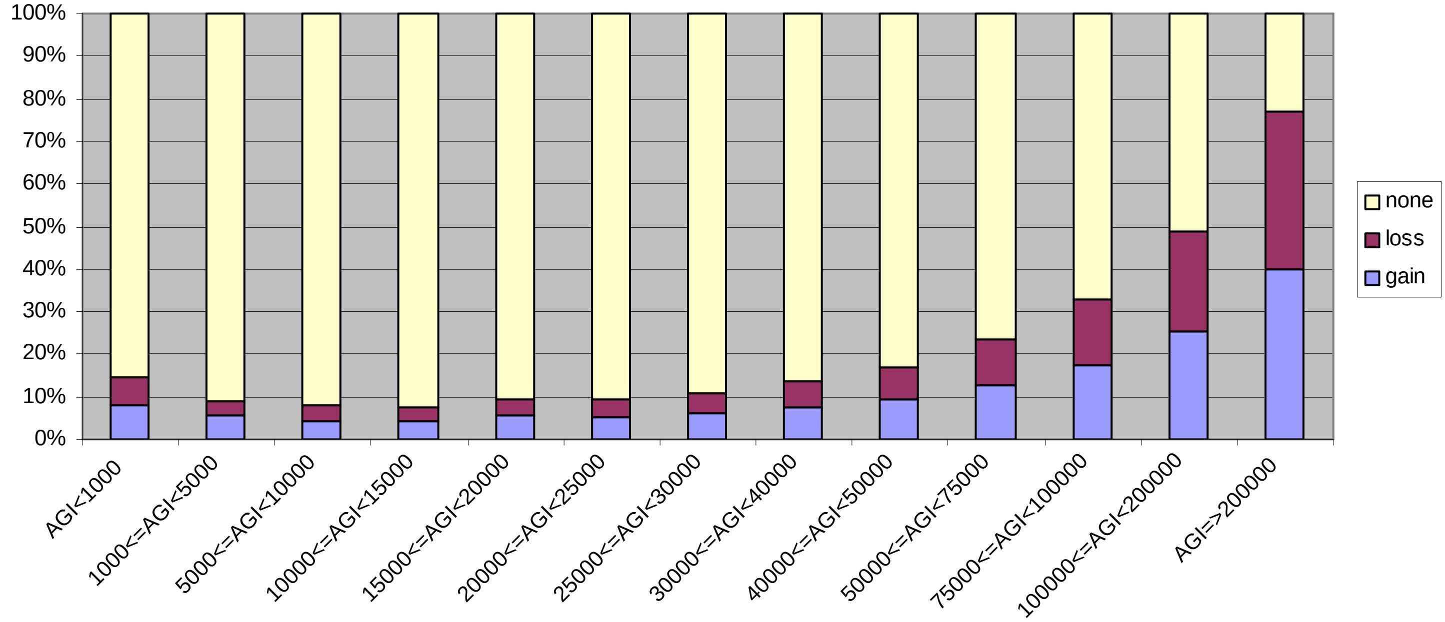
Figure 2: Aggregate Capital Gains by AGI Percentile



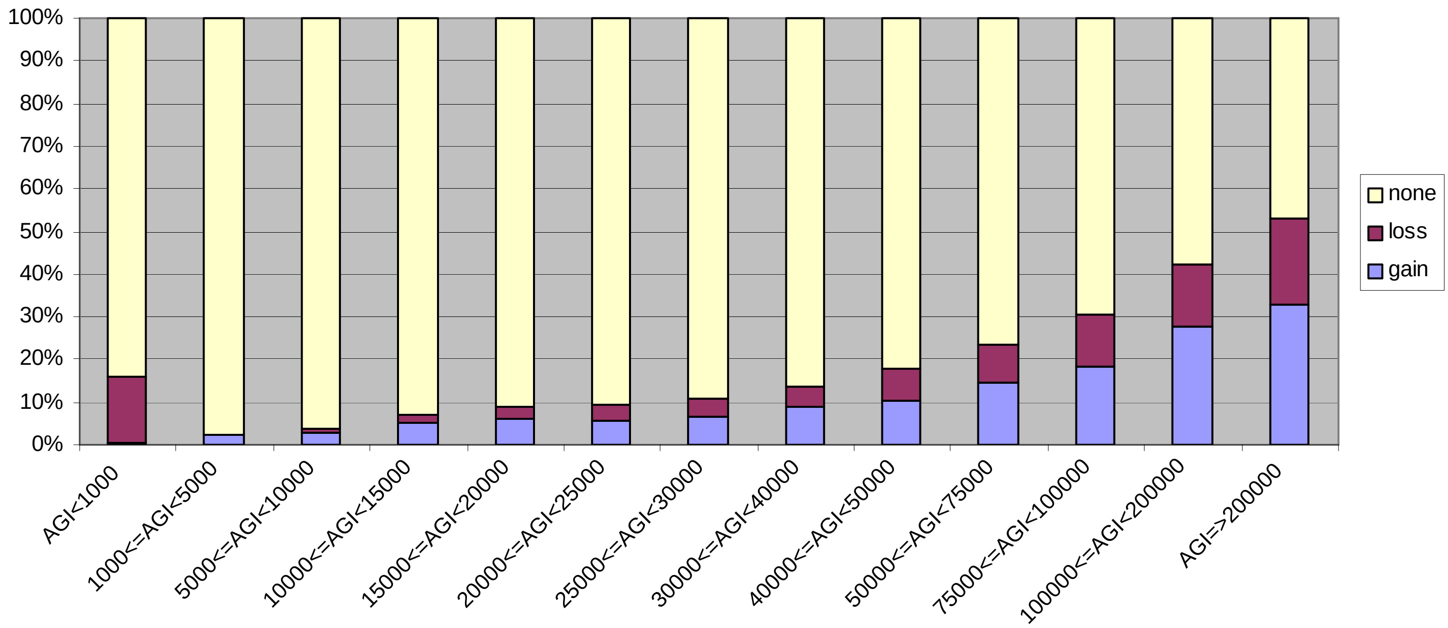
This chart shows the weighted aggregate amount of capital gains by AGI percentile for three years of the CPS ASEC tax model and two years of IRS Statistics of Income (SOI) public use microdata. The IRS file is used as the base data for the statistical match used to impute capital gains for tax years 2003 and 2004. Tax years 2002 and 2003 imputed gains to the lower 90% of the income distribution in alignment with the IRS but failed to assign sufficient quantities of capital gains to the highest income percentiles. Tax year 2004 (2005 ASEC) shows an appropriate increase in gains assigned to the top five percent of the income distribution.

Figure 3: Capital Gain and Loss Imputation

SOI TY01

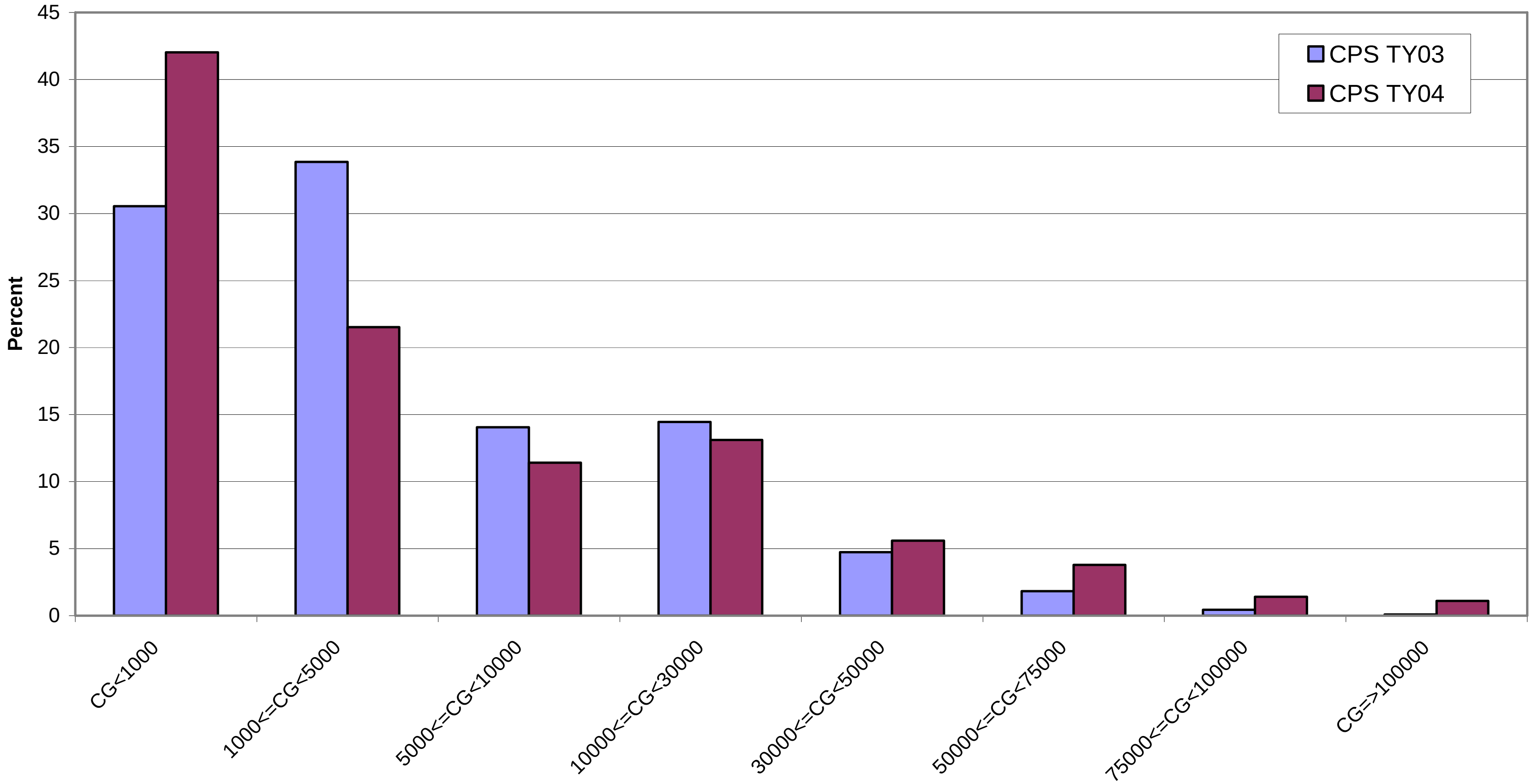


CPS TY04



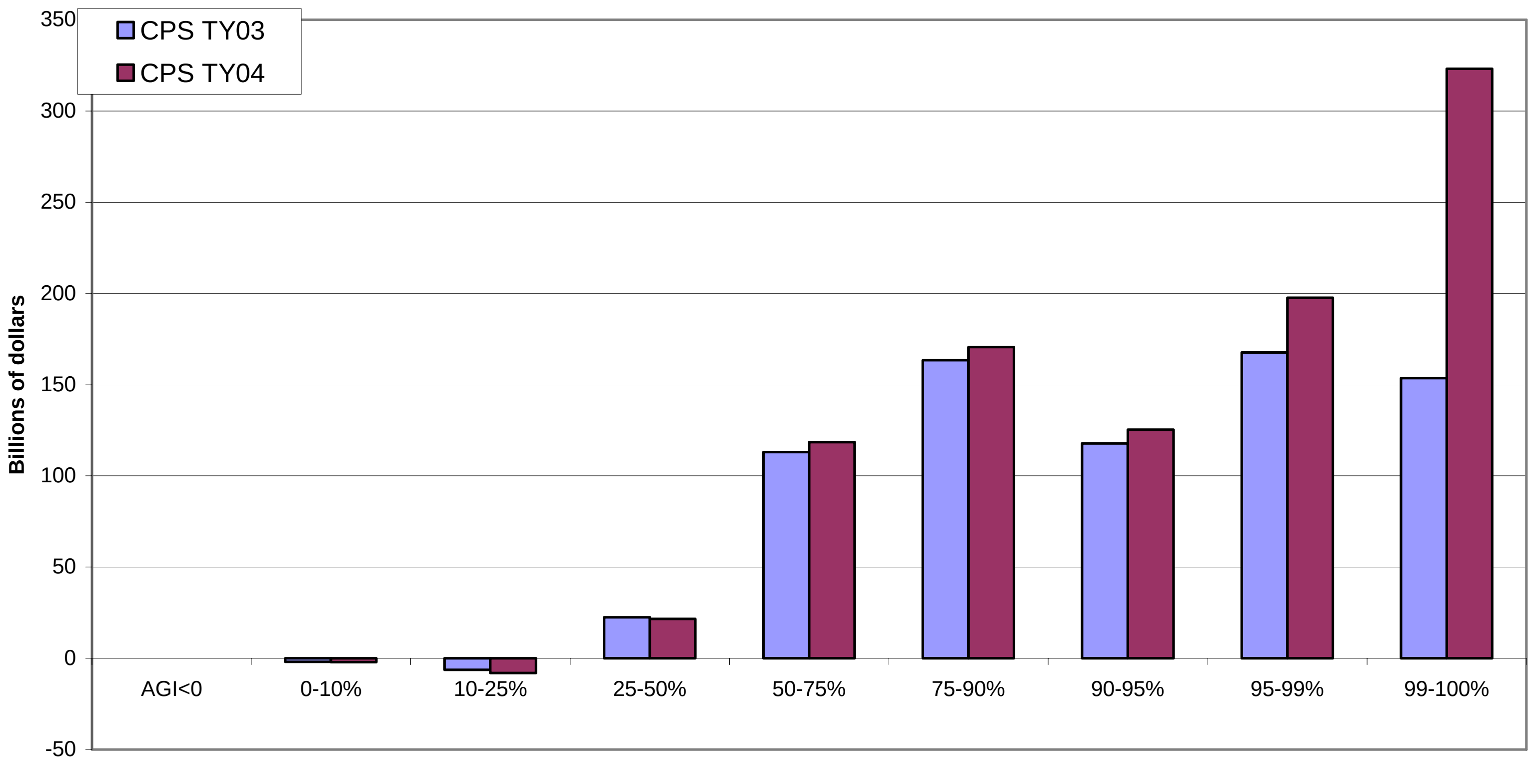
These charts show that capital gains assigned to CPS filers in the statistical match for tax year 2004 replicated the distribution from the source data (2001 Statistics of Income public use file). For each AGI class, the bar shows the percent assigned capital gains, capital losses or neither, summing to 100%. Capital gains incidence was constrained for the lowest AGI group this year using income-to-poverty ratios.

Figure 4: Capital Gains 2003-2004



This chart compares capital gains assigned in tax years 2003 and 2004 through the statistical match between the CPS ASEC and IRS Statistics of Income (SOI) public use file. The tax year 2004 match assigned more capital gains of less than $1000, and more over $50,000 compared to the imputed values for tax year 2003. The chart refers to all returns assigned capital gains and the amount categories sum to 100%. In 2004, 42% of files received less than $1000 in capital gains, compared to 30.5% of filers in 2003. More than five times as many filers were assigned $100,000 or more in capital gains in 2004 (2.5%) as in 2003 (0.4%).

Figure 5: Aggregate Federal Taxes After Credits by AGI Percentile



The weighted aggregate amount of federal taxes increased from $729 billion in 2003 to $947 billion in 2004. This chart shows that the increase focused on the upper end of the income distribution, particularly filers in the top one percent of AGI.

Table 1: State Tax Models[1]

State	Simulated items for Tax Year 2004
Alabama	Taxable Social Security exempt Unemployment compensation exempt Federal and military pensions excluded No self-employed tax deduction Different itemized deductions
Arizona	Pension exclusion Non-refundable family income tax credit
Arkansas	Spouse pension exemption separated Armed forces earnings exempt Unemployment compensation exempt Taxable Social Security exempt Non-refundable personal credit Non-refundable childcare expense credit
California	Taxable Social Security exempt Different itemized deductions Exemption credit Refundable childcare expense credit
Colorado	Non-taxable pensions excluded Refundable childcare expense credit
Connecticut	Adjusts taxable Social Security Non-refundable personal credit
Delaware	Pension exclusion based on age Aged-disabled exclusion Social Security exempt Split assets between spouses to apply pension exclusion Non-refundable personal credit Non-refundable childcare expense credit
District of Columbia	Social Security exempt Non-refundable low income credit Non-refundable childcare expense credit Refundable EITC[2] $3000 pension exclusion Disability exclusion
Georgia	Social Security exempt Retirement income exclusion Refundable low income credit
Hawaii	Social Security exempt Pensions exempt IRA exempt if over age 70 Refundable childcare expense credit Refundable low income credit Refundable low income renter credit
Idaho	Pension exclusion Social Security exempt Refundable grocery credit Non-refundable childcare expense credit

[1] Note that seven states (Alaska, Florida, Nevada, South Dakota, Texas, Washington, and Wyoming) are not included in the table because they do not levy individual state income taxes.
[2] Cannot claim both earned income credit and low income credit.

Illinois	Social Security exempt Excludes Armed Forces income Non-refundable EITC
Indiana	Exemption for aged with AGI below $40,000 Exemption for aged spouse Social Security exempt Disability exclusion Refundable unified credit for elderly Can deduct part of federal pension, military pay Part of unemployment is taxable Refundable EITC County credit
Iowa	Pension exclusion Adjusted Social Security exemption Armed forces income exempt Disability exempt Alternative tax for low income Non-refundable EITC Refundable childcare expense credit Non-refundable exemption credit
Kansas	Pension deduction Food sales tax refund Non-refundable childcare expense credit Refundable EITC
Kentucky	Separate income for spouses Non-refundable childcare expense credit Non-refundable low income credit Pension deduction
Louisiana	Social Security exempt Public pensions exempt Portion of dependent care expense credit allowed Childcare expense credit (refundable if low income)
Maine	Low income credit Non-refundable EITC Non-refundable elderly credit Social Security exempt Pension exclusion Partially refundable childcare expense credit
Maryland	Military pension excluded Disabled pension excluded Social Security exempt Dual income exclusion Partially refundable EITC Non-refundable poverty level credit Non-refundable childcare expense credit Local tax assumed to be 0.029%
Massachusetts	Part of FICA is deductible Assume 75% of income is not state generated Separate tax on capital gains and dividends Non-refundable limited income credit Refundable EITC Deduction for children and aged dependents

Michigan	Part of unearned income for aged deductible Military pay exempt Part of unemployment compensation excluded No self-employment tax deduction Social Security exempt Pension exclusion
Minnesota	Aged and disabled exclusion Married joint earners credit Refundable childcare expense credit Refundable EITC
Mississippi	Social Security exempt No self-employment tax deduction All pensions excluded
Missouri	Social Security exempt Pension exclusion
Montana	Unemployment compensation exempt No self-employment tax deduction No self-employment health insurance deduction No self-employment pension deduction Interest exclusion for aged Pension exclusion Military pay for active duty excluded Childcare expense credit
Nebraska	Standard deduction added back to high income returns Non-refundable personal exemption credit Non-refundable disabled/elderly credit Refundable childcare expense credit
New Hampshire	Only federal interest and dividends are taxed
New Jersey	Social Security exempt Unemployment compensation excluded Pension exclusion Retirement exclusion Refundable EITC
New Mexico	Low income rebate
New York	Social Security exempt Pension exclusion Disability pay exclusion limit Alternate tax for high income returns Non-refundable state household credit Non-refundable New York City household credit Refundable childcare expense credit Refundable EITC Refundable New York City school tax credit
North Carolina	Pension exclusion Refundable childcare expense credit Social Security exempt Refundable child tax credit Refundable elderly/disabled credit
Ohio	Non-refundable credit for elderly Social Security exempt Disability exclusion Non-refundable childcare credit Refundable joint filing credit Non-refundable retirement income credit

Oklahoma	Social Security exempt Pension exclusion Active military pay exclusion Non-refundable childcare expense credit Refundable EITC
Oregon	Social Security exempt Non-refundable retirement credit 60% Federal pension exclusion Military pay exclusion Non-refundable working family credit Non-refundable elderly/disabled credit Non-refundable EITC Non-refundable childcare expense credit
Pennsylvania	IRA distributions exclusion Pension exclusion Unemployment compensation exclusion Social Security exempt Refundable Taxback credit
Rhode Island	Non-refundable childcare expense credit Non-refundable dependent care expense credit Partially refundable EITC
South Carolina	Social Security exempt No disabled pension deduction Aged deduction Pension exclusion Extra exemption for children under 6 Non-refundable dual earner credit Non-refundable childcare expense credit
Tennessee	No tax if aged and meet AGI test
Utah	Federal tax deduction Retirement income exclusion Pension exclusion
Vermont	Non-refundable child and disability credits Refundable EITC
Virginia	Social Security exempt Age 62+ tax exclusion Age 65+ tax exclusion Childcare expense deduction Refundable low income credit
West Virginia	Military pension deduction Government pension exclusion Senior citizen or disabled deduction Low income earner exclusion
Wisconsin	Unemployment compensation partially deductible Married couple dual earner credit Working family credit Pensions deductible Disabled exclusion Itemized credit Refundable EITC

Table 2: Variables on 2005 ASEC Public Use File

FILESTAT	- same
DEP_STAT	- same
CTC_CRD	- new, child tax credit
EIT_CRED	- same
ACTC_CRD	- new, additional child tax credit
FICA	- same
FED_RET	- same
AGI	- same
CAP_GAIN	- same
CAP_LOSS	- same
TAX_INC	- same
MARG_TAX	- same
FEDTAX_BC	- new, federal taxes before refundable credits
FEDTAX_AC	- new, federal taxes after all credits
STATETAX_B	- new, state taxes before refundable credits
STATETAX_A	- new, state taxes after all credits

FED_TAX from previous years can be recreated by adding the earned income credit (EIC) to the after credit amount. In previous years, FED_TAX was the "total" amount of federal taxes before the EIC. It included child and additional child credits.

FED_TAX = FEDTAX_AC + EIT_CRED

STATETAX_A is the same as STATETAX from previous years, an amount that has been reduced by all credits. Prior to tax year 2003, state tax credits were not modeled as refundable. The new methodology does permit refundable credits, therefore before (refundable) credit and after all credit estimates are now being released.

Table 3: Weighted Aggregate Income, Credits and Tax

Variable	Aggregate Dollars CPS TY04[1]	Aggregate Dollars CPS TY03	Aggregate Dollars IRS TY03[2]	CPS TY04/ IRS TY03	CPS TY03/ IRS TY03
Adjusted Gross Income	6,386,284,231	6,182,497,186	6,212,974,067	102.8%	99.5%
Itemized Deductions	657,959,091	766,040,626	880,553,392	74.7%	87.0%
Taxable Income	4,626,005,684	4,341,022,379	4,202,698,688	110.1%	103.3%
Child Care Credit	3,454,255	2,990,893	3,222,404	107.2%	92.8%
Child Tax Credit	36,430,239	35,161,784	25,741,752	141.5%	136.6%
Additional Child Tax Credit	12,703,786	7,153,168	12,627,142	100.6%	56.6%
Earned Income Credit	26,312,121	25,280,285	39,124,190	67.3%	64.6%
Federal Taxes (after all credits)	947,051,632	729,767,186	750,024,250	126.3%	97.3%
	Number of Returns CPS TY04	Number of Returns CPS TY03	Number of Returns IRS TY03	CPS TY04/ IRS TY03	CPS TY03/ IRS TY03
Total Number of Returns	123,623,100	123,556,400	130,571,319	94.7%	94.6%
Itemized Deductions	40,156,290	42,732,300	43,861,971	91.6%	97.4%
Taxable Income	103,905,460	103,720,150	101,669,737	102.2%	102.0%
Child Care Credit	6,411,188	6,497,370	6,340,695	101.1%	102.5%
Child Tax Credit	26,364,800	25,928,400	22,815,814	115.6%	113.6%
Additional Child Tax Credit	11,306,480	7,219,220	9,131,488	123.8%	79.0%
Earned Income Credit	16,825,010	16,170,300	22,233,971	75.7%	72.7%
Federal Taxes (after all credits)	113,507,720	112,481,000	89,103,880	127.4%	126.2%

This table compares aggregate results from the CPS TY04 tax model with the most recent results published by the IRS, preliminary data for TY03. The top panel of the table compares weighted aggregate dollar amounts. The lower panel compares the number of returns for each variable between the two data sources. The model has lower amounts of itemized deductions and earned income credits assigned, and higher child tax credits and total tax amounts. Looking at the number of returns, the model assigns has fewer earned income credit recipients and more filers with non-zero federal tax values.

[1] Aggregate dollar values in thousands for both CPS and IRS.
[2] IRS data from Statistics of Income Bulletin (Volume 24, Number 3) Winter 2004-2005.