Methodology Appendix to Burman, Geissler, and Toder, "How Big are Total Individual Income Tax Expenditures and who Benefits from them?"

We used the Tax Policy Center (TPC) microsimulation model to estimate tax expenditures and distributional effects. The model calculates tax liabilities for a weighted sample of individual income tax returns and can be used to analyze current law as well as alternative policies that change the tax code, including tax rates, the AMT, exemptions, deductions, and credits. It is based on the 2001 Public Use File (PUF) produced by the Statistics of Income Division (SOI) of the IRS. The PUF includes sample weights that can be used to produce population estimates. The PUF and sample weights are extrapolated to later years based on more recent SOI published data and Congressional Budget Office (CBO) economic projections, and imputations from other data sources of items not reported on tax returns. The model assumes that, conditional on income, deductions, family status, etc., individuals legally minimize their tax owed (for example, by correctly deciding whether to itemize deductions), but it does not incorporate any other types of behavioral responses, such as changes in labor supply, investment or consumption patterns in response to tax incentives.

To analyze a policy option, the model performs two simulations – a baseline simulation and a simulation of the proposed. For this paper, we use two different baselines – 2007 law prior to the enactment of the AMT patch, and 2007 law with the AMT eliminated. The revenue and distributional estimates are generated by comparing the numbers from the two baselines with alternative simulations that assume separate tax expenditures and combinations of tax expenditures are eliminated.

Most of the data required for the simulations are reported on individual income tax returns in the PUF sample. Some simulations, however, required data not reported on individual tax returns and not previously imputed to the TPC model. For this paper, we imputed data on retirement account balances, life insurance interest, and health insurance coverage and premiums to individuals in the TPC model database. The methods used for these imputations are described below.

A full description of the TPC microsimulation model is in Rohaly, Carasso, and Saleem (2005).

<u>Imputing Defined Contribution, IRA, and Life Insurance Accruals</u>

While the retirement module in the TPC tax model already contains imputed information on pensions and contributions, it does not have any information on account balances. In order to treat pension accounts like other savings accounts and life insurance interest as income, we must include the accruals in these accounts in our calculation of income. To estimate the values of individuals' Defined Contribution (DC), IRA and Keogh accounts,

and life insurance interest, we use a similar methodology to that described in Appendix A of Rohaly, Carasso, and Saleem (2005).[1]

We run probit regressions on the probability of having a positive acccount balance using data from the Federal Reserve Board's Survey of Consumer Finances (SCF) for 2001. We use an exhaustive set of the independent variables that are included in both the SCF and the PUF. These variables include number of dependents, age (included as 10-year bracket dummies), income (as defined in the SCF), and the following components of income: income from a farm or business, tax-exempt interest income, taxable interest income, rental income from schedule E, pension income, taxable dividends, and realized capital gains (all defined as the natural logarithm of the income item plus one). We also include dummies for zero values of each income item; dummies for negative overall income, negative income from a business or farm, and negative capital income; as well as interactions between the negative income dummies and the appropriate negative income amount (defined as the natural logarithm of the absolute value of the income item plus one). In addition, we include dummies for whether the individual itemizes deductions on his or her federal tax return, and dummies for whether certain federal tax schedules are filed (C for business income, E for rental income, and F for farm income).[2] While most of these variables are given at the household level, age and wages are individual level variables. Conditional on the individual having a positive account balance, we regress the natural log of this account balance against this same set of independent variables to generate a second set of coefficients.[3]

After generating these two sets of coefficients, we apply them to the observations in the tax model. Applying the parameter estimates from the probit to data in the PUF, we compare the probability of having a positive account balance to a pseudo-random (uniformly distributed) number to impute account status. For those imputed to have an account, the level is based on the ordinary least squares estimates applied to the PUF data. We scale both the imputed probability indexes and levels in the TPC model so that our projections roughly match the SCF data.

For each type of account, we assume a 7 percent rate of return[4] and therefore estimate annual accruals to be equal to 7 percent of the value of the account. We treat this income as capital gains for DC, IRA, and Keogh accounts and as ordinary income for life insurance interest.

---

[1] One major difference is that our regressions are not censored from above as there is no limit to how much an individual can have in these accounts whereas the regressions for contribution levels are censored to account for annual contribution limits.

[2] This list and description of the independent variables is taken directly from Rohaly, Carasso, and Saleem (2005). With each retirement account, we run regressions for both the head and spouse. For life insurance, we do this calculation at the household level using head age and wage figures because the SCF only provides life insurance information at the household level.

[3] When estimating the Defined Contribution account balances, we also included imputed DC contributions to the account made this year as an independent variable.

[4] This is roughly equal to the rates of return found in CBO (2004) available at http://www.cbo.gov/ftpdocs/54xx/doc5418/05-10-RetirementSavings.pdf.

Imputing Defined Benefit Accruals

The retirement module of the tax model uses a similar methodology to that discussed above and in Rohaly, Carasso, and Saleem (2005) to determine whether the individual has a Defined Benefit (DB) plan, but does not have any information on annual accruals. We used the Urban Institute's DYNASIM model[5] to generate estimates of the DB plan accrual to earnings ratios by income quintile and age bracket conditional on the individual having a defined benefit plan. This value does not estimate contributions or plan earnings, but rather the accrual in the present value of future benefits. For some subgroups, this accrual value was negative[6] and when this occurred, we treated it as if there was no accrual. We then use the information in the tax model that assigns some filers DB plans. If an individual is assigned a plan, we multiply their wages by their subgroup's accrual to earnings ratio to estimate their DB accruals over the past year. We add this value to their adjusted gross income.

Imputing Health Insurance Coverage and Premiums[7]

Income tax returns do not include information on employer-sponsored health insurance. Thus, it must be imputed from other sources. We do this in several steps. First, we match estimates of employer-sponsored health insurance coverage, non-group health insurance coverage, insurance premiums, and employer and employee payment shares for employer-sponsored insurance based on data from the Urban Institute's Transfer Income Model (TRIM).[8] To do this, tax units in the TRIM and tax model databases are partitioned into cells based on adjusted gross income (AGI), age, filing status, and the presence and number of dependents. The overall prevalence of insurance and the distribution of values in the TRIM database are calculated for each cell, and these are used to assign values to tax units in the tax model database in the corresponding cell. This random assignment is carried out subject to two restrictions: tax units with head under age 55 may not receive employer-sponsored insurance if they do not have wages and tax units claiming the self-employed health insurance deduction must be assigned health insurance equal to the value of the deduction.

Imputed premium values are based on data from 2000 and 2001. Through 2006, the values are grown at the rate of premium growth reported in the 2006 Kaiser/HRET Employer Health Benefits Survey. Premiums are projected to grow at the rate of national medical expenditures per capita from the National Health Expenditure Accounts through

[5] For more on the DYNASIM model, see Favreault and Smith (2004).
[6] This only occurred in some cases for older age groups as likely stems from the fact that under some pension formulas, it is possible that another year of work actually reduces the present value of pension benefits. This occurs because the present value of the increased annual pension from working another year is less than the loss in present value from delaying pension benefit receipt by one year.
[7] This section is taken from Appendix 3 of Burman, Furman, Leiserson, and Williams (2007).
[8] For additional details on the TRIM database see the appendix to Burman, Uccello, Wheaton, and Kobes (2003).

2015 and at the 2015 rate in 2016 and 2017. Coverage rates and other variables are assumed to be unchanged over the period.

After the health insurance determination, tax units with coverage are assigned to family coverage or individual coverage based on the value of the plan they have and the distribution of plans by value from the 2006 Kaiser survey. Using estimates based roughly on the 2003 Kaiser survey, a percentage of tax units based on income class are designated as having access to premium conversion plans allowing them to pay health insurance premiums on a pre-tax basis. Finally, tax units are assigned to full-year or part-year coverage based on the prevalence of both types of coverage as reported in the Medical Expenditure Panel Survey.

Limitations. There is inherent uncertainty in our estimates—and any estimates—of future health insurance coverage and premiums. They require projections of health insurance premium growth, which may vary systematically by income and over time. They require highly uncertain estimates of coverage by premium conversion plans, for which few data are publicly available. In future work, we plan to base our estimates on more recent and complete information about health insurance premiums and coverage. For all of these reasons, our estimates may differ from published estimates by the Department of the Treasury and other sources, and may be significantly revised in the future. In addition, Treasury estimates account for behavioral responses such as people purchasing health insurance in response to the tax incentives. We do not yet have the capability of making such estimates.

Comparisons with OMB and JCT Estimates

Table 1 shows comparisons of our tax expenditure line estimates with JCT and OMB reported tax expenditure estimates. (Treasury's Office of Tax Analysis performs the estimates for OMB). In most cases, our estimates are similar to theirs, although in some cases either OMB or JCT combine different line items, so that the provisions they are estimating are not always strictly comparable either with our estimates or with each other's.

An exception is our estimate for the tax expenditure on the exclusion of tax-exempt interest, which is significantly smaller than those reported by the Joint Committee on Taxation (JCT) and the Office of Management and Budget (OMB). We use tax return data from the PUF to determine taxpayers' tax-exempt interest, whereas JCT includes trends in the bond market in their analysis. Johnson and Moore (2005) find that IRS and SCF estimates of nontaxable interest are similar, which leads us to believe that the tax return data used in the model provides a good estimate. Even if the true cost of this exclusion is closer to the values estimated by JCT and OMB, the focus of this paper is on the interactions between expenditures and we would not expect the difference between our estimate and the true value to affect these interactions much.

References

Burman, Len, Jason Furman, Greg Leiserson, and Roberton Williams (2007). "The President's Proposed Standard Deduction for Health Insurance: An Evaluation." Tax Policy Center at http://www.taxpolicycenter.org/UploadedPDF/411423_Presidents_Standard_Deduction.pdf.

Burman, Leonard E., Cori E. Uccello, Laura L. Wheaton, and Deborah Kobes (2003). "Tax Incentives for Health Insurance." Tax Policy Center Discussion Paper No. 12.

Congressional Budget Office (2004). "Tax-Deferred Retirement Savings in Long-Term Revenue Projections." Congressional Budget Office at http://www.cbo.gov/ftpdocs/54xx/doc5418/05-10-RetirementSavings.pdf.

Favreault, Melissa and Karen Smith (2004). "A Primer on the Dynamic Simulation of Income Model (DYNASIM3)." Urban Institute Discussion Paper 02-04.

Johnson, Barry and Kevin Moore (2005). "Consider the Source: Differences in Estimates of Income and Wealth From Survey and Tax Data." Federal Reserve Board Working Paper.

The Kaiser Family Foundation and Health Research and Educational Trust (2006). "Employer Health Benefits: 2006 Annual Survey." At http://www.kff.org/insurance/7527/upload/7527.pdf.

Rohaly, Jeffrey, Adam Carasso, and Mohammed Adeel Saleem (2005). "The Urban-Brookings Tax Policy Center Microsimulation Model: Documentation and Methodology for Version 0304." Tax Policy Center at http://www.taxpolicycenter.org/UploadedPDF/411136_documentation.pdf.

Table 1. Selected Individual Income Tax Expenditures, Individually and in Groups, Tax Year 2007

	TPC Estimates		OMB	JCT	AMT
	w/ AMT	**no AMT**	(FY07)	(FY07)	**Taxpayers**
	(billions of dollars)				**(millions)**
Exclusion of interest on life insurance savings	18.3	17.5	18.3	26.1	24.0
Net exclusion of contributions and earnings for retirement plans[1]	126.8	119.9			24.3
Employer contributions			49.2	108.6	
Employee contributions to DC plans			42.4		
IRAs			5.7	15.5	
Keogh plans			10.9	8.8	
Exclusion of interest on tax-exempt bonds	12.2	11.6	26.2	25.4	23.6
Exclusion of employer contributions for medical insurance	137.7	129.5	141.3	105.7	28.4
Exclusion of Social Security and railroad retirement benefits	23.0	22.6	26.9	22.4	23.9
Exclusion of veterans benefits	5.2	5.0	4.2	3.5	23.6
Subtotal: Exclusions from Income, without interactions	**323.3**	**306.0**	325.1	316.0	
Subtotal: Exclusions from Income, with interactions	**343.9**	**325.5**			29.7
percentage change	6.4%	6.4%			
Deductibility of Student Loan Interest	1.1	1.0	0.8	0.9	23.5
Self-employed medical insurance premiums	3.8	3.7	4.4	3.8	23.5
Additional deduction for the blind and elderly	1.5	1.9	2.2	1.7	23.3
Subtotal: above the line deductions, without interactions	**6.4**	**6.6**	7.4	6.4	
Subtotal: above the line deductions, with interactions	**6.4**	**6.6**			23.4
percentage change	0.0%	0.1%			
Lower tax rates on long-term capital gains	83.7	86.4	53.1	127.1	23.8
Lower tax rates on qualifying dividends	11.1	10.9	0.0		23.7
Subtotal: special tax rates, without interactions[2]	**94.8**	**97.3**	53.1	127.1	
Subtotal: special tax rates, with interactions	**96.0**	**99.0**			23.9
percentage change	1.2%	1.8%			

Deductibility of Mortgage interest on owner-occupied homes	92.4	79.9	79.9	73.7	28.8
Deductibility of State and Local Taxes	39.0	76.5			19.4
Property taxes on residences			15.5	16.8	
Income and other taxes			33.7	33.9	
Deductibility of Charitable Contributions	43.3	38.5	47.4	41.9	24.9
Deductibility of Casualty Losses	0.4	0.3	0.3	0.8	23.5
Deductibility of Medical Expenses	5.6	5.4	4.2	8.4	23.6
Subtotal: itemized deductions, without interactions	**180.7**	**200.6**	181.0	175.5	
Subtotal: itemized deductions, with interactions	**153.5**	**153.2**			29.4
percentage change	-15.0%	-23.6%			
HOPE tax credit	2.6	3.7	3.3	3.1	23.1
Lifetime learning tax credit?	1.6	2.2	2.2		23.2
Credit for child and dependent care expenses[3]	1.9	3.4	2.8	3.0	23.1
Low and moderate income savers credit	1.9	1.9	0.7	0.9	23.5
Subtotal: non-refundable credits, without interactions	**8.0**	**11.1**	9.0	7.0	
Subtotal: non-refundable credits, with interactions	**8.2**	**11.3**			22.5
percentage change	2.8%	1.6%			
Child credit[4]	44.9	44.9	47.5	45.0	23.6
Earned income tax credit[4]	43.7	43.7	41.8	44.7	23.5
Subtotal: refundable credits, without interactions	**88.6**	**88.5**	89.3	89.7	
Subtotal: refundable credits, with interactions	**89.2**	**89.1**			23.6
percentage change	0.6%	0.6%			
Total: all provisions without interactions	**701.8**	**710.2**			
Total: all provisions with interactions	**760.5**	**746.7**			35.4
percentage change	8.4%	5.1%			
Total: all provisions with interactions and revenue neutral tax cut	0.3	0.3			35.4
Addendum: Baseline individual income tax revenues, AMT taxpayers	1,020.7	950.9			23.5

Notes

1. Tax expenditure is revenue loss attributable to deduction/exclusion for contributions and earnings net of any tax on withdrawals
2. OMB does not consider the lower tax rates on capital gains and dividends on corporate stock to be a tax expenditure
3. JCT includes the value of the exclusion of employer provided child care
4. Includes both refundable and nonrefundable portion